

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2004, 21:40:18 ; Search time 5813 Seconds

(without alignments)
10766.788 Million cell updates/sec

Title: US-09-868-744B-1

Perfect score: 1444
Sequence: 1 atcatgattcctaatacatc.....gtactcgaatgtagatcc 1444

Scoring table: IDENTITY NUC
Gapop 10_0, Gapept 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_stg:*
12: gb_sy:*
13: gb_un:*
14: gb_vt:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pac:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_stg:*
28: em_un:*
29: em_vt:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pun:*
35: em_hcg_rod:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_higo_hum:*
40: em_higo_mus:*
41: em_higo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1442	99.9	1444	6	AX026532
2	1349.6	93.5	75508	8	AB026654
3	1341.8	92.9	6385	6	AX463618
4	1341.4	92.9	4526	6	AX463617
5	1306.2	90.5	11127	6	AX392814
6	1298.4	89.9	3450	6	AX463620
7	1297.4	89.8	3408	6	AX463619
8	1297	89.8	2857	6	AX463621
9	1234.2	85.5	3172	8	ATU41998
10	1184	82.0	1217	6	AX461227
11	1174.8	81.4	1219	6	AR236191
12	1174.8	81.4	1219	6	AR438153
13	1174.8	81.4	1219	6	AX167634
14	1174	81.3	1742	6	AR236212
15	1174	81.3	1742	6	AR438174
16	1174	81.3	1742	6	AX167655
17	1078.8	74.7	15676	6	BD251968
18	1078.8	74.7	17111	6	BD251969
19	1078.8	74.7	17116	6	BD251970
20	475.8	33.0	11461	6	BD266633
21	475.8	33.0	11461	6	AX047314
22	474.8	32.9	12766	6	BD266630
23	474.8	32.9	12766	6	AX047311
24	163	11.3	1486	8	AX087751
25	158	10.9	1679	8	AY064043
26	158	10.9	2637	8	AY056800
27	153	10.6	1559	8	AF428330
28	96.8	6.7	119091	8	AC016041
29	95	6.6	122	8	ATU524576
30	94.4	6.5	3261	8	ATU42007
31	76	5.3	2000	6	AX508607
32	75.4	5.2	1271	6	AR236192
33	75.4	5.2	1271	6	AR438154
34	75.4	5.2	1271	6	AX167635
35	73.8	5.1	1800	6	AR236211
36	73.8	5.1	1800	6	AR438173
37	73.8	5.1	1800	6	AX167654
38	62.6	4.3	2460	8	STPOAC71
39	62.2	4.3	253001	3	AE014834
40	61.6	4.3	251448	3	AE014819
41	59.4	4.1	349751	3	PFMAL4P3
42	59.2	4.1	128683	2	AC135793
43	59	4.1	11861	9	AC069435
44	58.4	4.0	1809	1	AP267204
45	57	3.9	1809	1	AP267216

ALIGNMENTS

RESULT 1
AX026532
LOCUS AX026532 1444 bp DNA linear PAT 16-SEP-2000
DEFINITION Sequence 1 from Patent WO0037661.
ACCESSION AX026532
VERSION AX026532.1 GI:10187719
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE
Pettig, J.H., Schepers, F.M. and van Dun Cornelis, M.P.
AUTHORS
TITLE Genetic modification of compositae

'product' or 'note' qualifiers. Genes that have no significant protein similarity are described as 'unknown protein'. The software programs used to predict genes include: Graal (Informatics Group, Oak Ridge National Laboratory, <http://compbio.ornl.gov/Graal-1.3/>), GENSCAN (Chris Burge, MIT, <http://CCR-081.mit.edu/GENSCAN.html>), NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of Denmark, <http://www.cbs.dtu.dk/services/NetGene2/>) and SplicePredictor (Volker Brendel, Stanford University, <http://gremiin.zool.iastate.edu/cgi-bin/sp.cgi>). Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, Washington University School of Medicine, St. Louis, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). This sequence may not be the entire insert of this clone. It may be shorter because we remove overlaps between neighboring submissions. The 5' clone is KX499 and the 3' clone is MCB22.

FEATURES

source

1. 75508

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/strain="Columbia"

/db_xref="taxon:3702"

/chromosome="3"

/clone="WVE11"

/clone_lib="Mitsui P1"

complement(1178..2338)

/note="gene_id:MVE11.1"

unknown protein"

/codon_start=1

/evidence=not experimental

/protein_id="BAB01791.1"

/db_xref="GI:9293888"

/translation="MASSSSSSLSFSTSKNNKTPFKKNSAFSSRSATSLIKRQTFV
KKAKEISLTCIDIDVCYCYGNSAEKLTPEPEREKYALIAKRGELSEKRRKSVLH
EFLERKAKDDPEKSKKIKVRKRVKXPVMDRPNFNTSVQLMGVQSLEKRLRI
QHRTCVAVAGQQRVQYITMANQELMANNTQLOSHNSVMTLMNMGAFSQIP
VSLASNTQSLAPIPPELMIYNSDANYSGLGVGTGINGLOMNMMLYNNINSV
NDFSXQPDNSRAESYSLGVHEDGNNEFNPMSRNKNFVQDCAGLGMQAGCTN
GLOSMMDHYSNNNSINSGLSHOYQFPYNSQODRVPNLDOGNNTSL"complement(join(3045..3317,3774..4546,4652..5271,
5392..5628,5717..5793))

/note="gb|AAC34345.1"

gene_id:MVE11.2

strong similarity to unknown protein"

/codon_start=1

/evidence=not experimental

/protein_id="BAB01792.1"

/db_xref="GI:9293889"

/translation="MANSPPAPAPPTTNGDSRRRLSASIEALCKRRFRNSKGGRS
DMVPPNIINFTQDNKSSCCCTFKQIVKLLPILSATLFTIYSPEAHSLSHS
SSRWIRQDPYFEDLDINMDVTKLENIIEGRTIGVNFPSNIEORWEYSKSD
NGDERKVVNLNDVADKNVTMDALYPMWIDEBQETVPVCPNIPNIXIPRRLDILV
KLPRKRWMSRQVGRHLQLAAATVYASAKGTFRGVFSRCPINLFRCKDLYS
RSGVWLYKPNLDLRDLQPVGSCSLPLGIGQDRPSLGNKREAVATLLSHAVY
VCGAIAAQSIRQSGSTRDVLIVDNISGVRSGLEAAGMOIRTIORIRNPAEKDA
YENWNSYKRWLQWLDYDKIIFIDADLLILNIDPLFSMEPSAIGANNGLFNSGAV
IEPCNCTPOLMEHINIESNGDQGLNEFTWMAHRIKPMNPLKHPWIGEDDAK
BKTELFGAEPPLVYLHYGKPKMICYRDYDCNPSNDIPEVPAIDIAHRKMWVDA
MPQELHFCYLSKQAKLEYDRQAEANVADHMKIKRQPRKICIDLCNWKMSM
LRHWGSMNTDYESVTPPPIITVDRKSSLPQHNL"complement(join(7969..8583,8684..9123,9207..9332,
9437..9586,9675..10305))

/note="emb|CAF70981.1"

gene_id:MVE11.3

similar to unknown protein"

/codon_start=1

/evidence=not experimental

/protein_id="BAB01793.1"

/db_xref="GI:9293890"

/translation="MDREKGVPRGASRLAPERTGSPFRPSMOIDEADLOKDEIR
QENSTYVLPKUIDSGELATKDFLDNRNLTALINSNDPPIKAVLSGHITVIEE
IIRIHDPQVYKINDNGTATLYAATGIVYRAECUNKCGALYSVRANKKEIPIV
VASLYGKRLVOYLYSHTPLSDLDPCDSDDEKGNAGMLVTNCTVDGLYCLALDLIO

CDS

CDS

RYPLAVTRDSNDTALMALAQTPYAPSGRLAFWQRMVYSCIHIEKINPHEVNH
HHHQRNSQDINSIOEKMKHLYFPPIIRRYVYKLGHAQAKELDCCQIPPE
DAAOQKAGLNOALFKAVENGIVYEEMMGHDPDIWYKNSGNTIFFAVSROBK
IFSLIRYKAKKNTLNTMIFINMMLHAAVAPARALITGALLOQSELOMFE
VEKIVPQRKRWNLKQKTPKALFTDQKDVQGGKMKRATSCVVALITTM
FSSAFVPGVYSDMPYTHQRFKIFLIDBALISCSMSLMLFGLISGRFRED
FLSLPTKIVGLALFESMATMIVFVTLTGVSKLSWASQFMFLAVIPGMEV
VQPLVILFIPATYCCNVDPKRRVFKL"

complement(join(11367..11581,11799..11940,12035..12088,12453..12544,
12656..12882,12966..13142,13242..13354)

/note="contains similarity to uridylylate kinase"

gene_id:MVE11.4"

/codon_start=1

/evidence=not experimental

/protein_id="BAB01794.1"

/db_xref="GI:9293891"

/translation="MALPLVTCSPISYSSISRTSFVPLTRKRTFSNYSRY
LISCSLSLSDNGSSPDPMNGNGNGSLNGQSSFRLLSPGTSKPLKRRVLLK
VSGBALAGDEQNIIDPKVTMAIARVAVATVRIEVAIVVGGNIIRGSAVAGSGLD
RSSADYIGMLVYVMAIPLQATMESIGIPTRYQTAFFMSVAPYIRRAIRHLEKQ
VVIIPAGTGNPFPTDTAALRCAEINAEVVIKATNYDGVDDPKRPNARLLDSL
YQVTSKDSLVDMDTAITLCOENKIPVVPVNLSEGNIAAINGEVGTLIGTWNBI
VTTS"

complement(13589..14257)

/note="gb|AAF16598.1"

gene_id:MVE11.5

similar to unknown protein"

/codon_start=1

/evidence=not experimental

/protein_id="BAB01795.1"

/db_xref="GI:9293892"

/translation="MDPEYFAGNPNPSDQONQKOLQICGPRPSPLSVHKSHTKRP
PKHAPPPRPDQPPYIPREPVYIYAVSKVYVATSEFNWVYVLTGISGVFLSEG
GGGVSAARLSTENASPRGKEPARAPETVINTMEEAARAGGVAGTILSPSL
LPSTSTISPMTHQGMFSPALPLGSRFPAGFMSPPRSQGTSLVASTPADPFSHI
WDQD"complement(join(15556..15731,16089..16154,16246..16333,
16599..16739))

/note="gene_id:MVE11.6"

unknown protein"

/codon_start=1

/evidence=not experimental

/protein_id="BAB01796.1"

/db_xref="GI:9293893"

/translation="MPTSYSFSNPTNNNMFLENIFFSSLLVIALRFPDQVDRY
SYKMGKPLENAPLYKKOIGLAKKNIKNMVEKAVLENDIGLKKNIKPEVGEY
PCSPPLAPPTYASSNSDMDAPRHVQVSNVPTLNDSTGVSNGPWSHLL"

complement(18068..18315)

/note="gene_id:MVE11.7"

unknown protein"

/codon_start=1

/evidence=not experimental

/protein_id="BAB01797.1"

/db_xref="GI:9293894"

/translation="WGRDETITVYTPSEFFKCPISLDVMSPVSLCTGYTPRASIOR
WLDGNNTCPATMOLKTKFPVPLTLQRLINWSDSIGRHNGDSPVNLPPSGEVP
TKEDVNTATLMSLEMLKTVREVDOSNRNEFLSKMEFVPLVDIIRTKKTEL
VYAIRFLDSIKVREERLSNMLANDGCLTALILAIQGNLSKESVRLTMDISF
DAKSKIMARQGVLEWMSISITSSPSLISASLSPLTISKRVSRSLAAKA
ITKIKDLTERTVNAVTEKSLKLETSRSGRGLTGGDNGCVAGVTKLLVY
STVTERAAVYLMCLCYVREDKTVETERNNGTLLVVIYIYSSNSAMTROMAKDLI
KVLFNSSALAYETKTHMP"

complement(join(28267..29317,29494..29585))

/note="gb|AAD43172.1"

gene_id:MVE11.8

similar to unknown protein"

/codon_start=1

/evidence=not experimental

/protein_id="BAB01798.1"

/db_xref="GI:9293895"

/translation="MRTROTYPPIAESLTARSIVQALPASATISGNGPKKKKCNV
RGIMDKOIPDLQGLILSRIGLKNHLSIVKTLKAVSVKFGSRLYPPQSQR
GGPKEGVYVLPNRSQTHHKKPELTGRNKLCAKQGMVLVYKNDPVYVFLNPF

CDS

CDS		GERICLPQVPQNSRDCLTSAAPTSTCCVISTFQSELYAVVVDWTRPGSVWTT HHFDKRRGEVFNRCISNGSFYCLSTSRGSLVDFSRVETWVLPVKECAPRRKML VRVFTHEGDIIVVTRVNNRKLAFKLNQGVNWEKMGVNLTVFSSDLSLT RAGLPEERNILYSSDIDDDFKVSHPTFVYVDSAMLPQPHDNFNF" complement (join (29896, 30389, 30465, .30582, 30745, .30879, 31009, 31129, 31210, .31276, 31364, .31468, 31677, .31788, 31869, .31981, 32106, .32223, 32374, .32474, 32694, .32810, 33023, .33152, 33271, .34000, 34112, .34210, 34701, .34792, 35000, .35176, 35688, .36488, 36613, .36879)) /note="gene id: MVE11.9 unknown protein" /codon_start=1 /evidence=not experimental /protein_id="BA01799.1" /db_xref="GI:923896" /translation="MGRDLVAAAKRAVRKAEVDRREQARWANNVGDILKNHGYVD ALKNFRIDYISVYKLPGLDLPCTQSLGRVILRLNFEELALYQKHGLQABEANT VEKQACTQLGRVTHBMFLKSEDDCEALQSAKKYFKAMELAQILKEKPPGESSGFL EYVNAHNNTGLMDLDDLPNAARTILKGLQICDEBEVREYDAARSRLHNLGNVFM ALRSNDRAKHLIEMDINICHKINHVQGEAGYINLAELHNKTKYQIDALLCYKASSL AKSQDSSALVQIENHTKI VKKSMYMEELREELMLKLSAEMTDAGTSEERKSM IQVACUGSLIDKSVMFAWLKHLQYSKRKKISDELCKEKLSDAFMLVBSQNRK NFRSLKWFIRSTVEGHEALGNLEGOALAKINGLDLCIGWTCALQYEBGYRIALK ANLPSIQLSALEDIHVTHMRFENAKASELKETIQNLKSEHAERKAECCSTQDCST DSEGHANVSNDRENACSSPQPNLSERLALDDEANDVPLISFLQPKRLFRKQV
Query Match		93.5%; Score 1349.6; DB 8; Length 75508; Best Local Similarity 98.7%; Pred. No. 3.4e-283; Matches 1423; Conservative 0; Mismatches 11; Indels 8; Gaps 6;
QY	1	ATTATGATCTCAATATGATATATATCTCATCTAGATCTAGGTTATCATATGTAAG 60
DB	57986	ATTATGATCTCAATATGATATATATCTCATCTAGATCTAGGTTATCATATGTAAG 58045
QY	61	AAAGTTTTCAGATATGNNACGACAAATGGCTACACTCGATCTGATTTGGTATCTCAAC 120
DB	58046	AAAGTTTTCAGATATGNNACGACAAATGGCTAGATCTGATTTGGTATCTCAAC 58105
QY	121	TCACATATATCTATACCAACATATAGTTAG-CAAAATTTAAACAACATAATTTTATGT 178
DB	58106	TCACATATATCTATACCAACATATAGTTAGCAAAATTTAAACAACATAATTTTATGT 58165
QY	179	ATGCAAGAGTCAGCATATGATATATGATTCAGATCGTTTTCAGAGTTCGGATGATG 238
DB	58166	ATGCAAGAGTCAGCATATGATATATGATTCAGATCGTTTTCAGAGTTCGGATGATG 58225
QY	239	AGTAGCCATTTATATGATCACTAATCGTGAATAGTG-ATAATGATCAACAATCTGAT 297
DB	58226	AGTAGCCATTTATATGATCACTAATCGTGAATAGTGATAATGATCAACAATCTGAT 58285
QY	298	CTTATTTGATATATATCCATATACATCATGAAAGACATTTCTTTTCAGGCTCGAAT 357
DB	58286	CTTATTTGATATATATCCATATACATCATGAAAGACATTTCTTTTCAGGCTCGAAT 58345
QY	358	AATTATGATACAAATCTAATAGAAAACGAAATTAATACGTTGAAATGTAAGAACTTAA 417
DB	58346	AATTATGATACAAATCTAATAGAAAACGAAATTAATACGTTGAAATGTAAGAACTTAA 58405
QY	418	TTGACACGCCAACACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 477
DB	58406	TTGACACGCCAACACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 58465
QY	478	CACAAAAAACCGAGTGTGATGTAACACCGGATCGAGAGGTCACAGTCATGAAGCC 537
DB	58466	CACAAAAAACCGAGTGTGATGTAACACCGGATCGAGAGGTCACAGTCATGAAGCC 58525
QY	538	ATCAAGCAAAAGAACTAATCCAAAGGGTGAGATGTAATAGTTTAAAAATAGTTAA 597
DB	58526	ATCAAGCAAAAGAACTAATCCAAAGGGTGAGATGTAATAGTTTAAAAATAGTTAA 58585
QY	598	CACGAGGAAA--CTGTCTGACACGAGTTCAGTTATCTTTTACCTGTGTCGAAATGA 656
DB	58586	CACGAGGAAAAGGCTGTCTGACACGAGTTCAGTTATCTTTTACCTGTGTCGAAATGA 58645

QY	657	TTCTGTCTGTGCGATTTAACTATATTTTTTTGAAAGCCGAAATAAAGTTGTAAAGAGATA 716
DB	58646	TTCTGTCTGTGCGATTTAACTATATTTTTTTGAAAGCCGAAATAAAGTTGTAAAGAGATA 58705
QY	717	AACCGCGCTATATAAATTCAATATATTTCTCCCGCTTTTGAATGCTCTCGTTGCTCTCC 776
DB	58706	AACCGCGCTATATAAATTCAATATATTTCTCTCCGCTTTTGAATGCTCTCGTTGCTCTCC 58765
QY	777	TCACATTTTCATCAGCGCTTTTGATCTCCCGCGACTTGACAGAGAAACAAGAGAGAGA 836
DB	58766	TCACATTTTCATCAGCGCTTTTGATCTCCCGCGACTTGACAGAGAGAAACAAGAGAGAGA 58825
QY	837	CTAAGAGAGAAAGTAAGAGATAATCCAGAGAGATTCATCTCCGTTTGAATCTCTCTCAA 896
DB	58826	CTAAGAGAGAAAGTAAGAGATAATCCAGAGAGATTCATCTCCGTTTGAATCTCTCTCAA 58885
QY	897	TCTCATCTCTCTCTCCGCTCTTTCTTTTCCAAAGTAATAGAACTTTCTGGATCTACTTT 956
DB	58886	TCTCA--TCTTCTCCGCTCTTTCTTTTCCAAAGTAATAGAACTTTCTGGATCTACTTT 58942
QY	957	ATTGCTGTGATCTCGATCTCTTTCTCTCAATCTCTTCCAGAGTAATAGAACTTTCTGGATCTACTTT 1016
DB	58943	ATTGCTGTGATCTCGATCTCTTTCTCTCAATCTCTTCCAGAGTAATAGAACTTTCTGGATCTACTTT 59002
QY	1017	GGATCTGTGAACCTCCACTAAATCTTTTGGTTTCTAGTAATCGATCTAGTTGACCGAT 1076
DB	59003	GGATCTGTGAACCTCCACTAAATCTTTTGGTTTCTAGTAATCGATCTAGTTGACCGAT 59062
QY	1077	CAGTTAGCTCGAATTATAGCTTACAGAAATTTGGCTTGACCTTGATGGAGAGATCCATGTT 1136
DB	59063	CAGTTAGCTCGAATTATAGCTTACAGAAATTTGGCTTGACCTTGATGGAGAGATCCATGTT 59122
QY	1137	ATGTTACCTGCGAAATGATTTCTATATGTGAATGAAATCTGAACTGTGAAGTAGATT 1196
DB	59123	ATGTTACCTGCGAAATGATTTCTATATGTGAATGAAATCTGAACTGTGAAGTAGATT 59182
QY	1197	GAACTGTGAACACTGTCAATTTAGATTGAATCTGAACTGTGAAGTAGATT 1255
DB	59183	GAACTGTGAACACTGTCAATTTAGATTGAATCTGAACTGTGAAGTAGATT 59242
QY	1256	TGTGTATAGATTCTTCCGAAACCTTAGGATTTGATGTCGATCGTTGAAACAGAGCTAT 1315
DB	59243	TGTGTATAGATTCTTCCGAAACCTTAGGATTTGATGTCGATCGTTGAAACAGAGCTAT 59302
QY	1316	TTCTGATCAATCAGGCTTTATTTGACTGTATTTGAACCTTTTGTGTTGCTGAGCTCA 1375
DB	59303	TTCTGATCAATCAGGCTTTATTTGACTGTATTTGAACCTTTTGTGTTGCTGAGCTCA 59362
QY	1376	TAAAAATGGCTGAGGCTGACGATTAATCAACCAATCGTGTGACAACTGTAAGTGAATG 1435
DB	59363	TAAAAATGGCTGAGGCTGACGATTAATCAACCAATCGTGTGACAACTGTAAGTGAATG 59422
QY	1436	GT 1437
DB	59423	GT 59424

RESULT 3
AX463618
LOCUS
AX463618
DEFINITION
Sequence 15 from Patent WO0248335.
ACCESSION
AX463618
VERSION
AX463618.1
GI:21886378
KEYWORDS
synthetic construct
synthetic construct
artificial sequences.
ORGANISM
Meagher, R.B. and Li, Y.
Metal resistant plants and phytoremediation of environmental
contamination
REFERENCE
Patent: WO 0248335-A 15 20-JUN-2002;
JOURNAL

1097 AGAGAGAAAGTAAGAGATAATCCAGAGATTCAATCTCCGTTTGAATCTTCCCTCAATCT 1156

q

17 AAAGTTTTCAGCAATATGNNACGACAAAATGGCTTACACTCGATCTCTAATTCATATCCAC 375

QY 121 TCACATATTACTATACCAACATTAGTTAGCAAAATTTAAACAACATATTTTATGTAT 180
Db 377 TCACATATTACTTATACCAACATTTAGTTAGCAAAATTTAAACAACATATTTTATGTAT 436
QY 181 GCAAGAGTCAGCATATGATTAATTTGATTCAGAAATCGTTTTCGAGAGTTCGGATGTAGTAG 240
Db 437 GCAAGAGTCAGCATATGATTAATTTGATTCAGAAATCGTTTTCGAGAGTTCGGATGTAGTAG 496
QY 241 TAGGCATTTATTTAATGTCATACACTAATTCGTTGATGTAGTATGATGAACATTTGATCTT 300
Db 497 TAGGCATTTATTTAATGTCATACACTAATTCGTTGATGTAGTATGATGAACATTTGATCTT 556
QY 301 ATTGTATAAATATCCATAAACAATCATGAAGAGACATTTCTTCAGGTCCTGAATTAAT 360
Db 557 ATTGTATAAATATCCATAAACAATCATGAAGAGACATTTCTTCAGGTCCTGAATTAAT 616
QY 361 TATGATACAAATTCATTAAGAAACGAATTAATTTACGTTGAATTTGATGAATCTAATTTG 420
Db 617 TATGATACAAATTCATTAAGAAACGAATTAATTTACGTTGAATTTGATGAATCTAATTTG 676
QY 421 AACAGGCCAACACGAGGAGTAACTGCTGCTGATTCGATTCGCTGTTTAACTTAACAC 480
Db 677 AACAGGCCAACACGAGGAGTAACTGCTGCTGATTCGATTCGCTGTTTAACTTAACAC 736
QY 481 TAAAAAACGAGGTGTCTATGTAACACGGGATCGAGCGTCAAGTCATGAAGCCATC 540
Db 737 TAAAAAACGAGGTGTCTATGTAACACGGGATCGAGCGTCAAGTCATGAAGCCATC 796
QY 541 AAAGCAAAAGAACTAATCCAGGGTGAGATTAATTTAGTTTAAAAATTTAGTTAACAC 600
Db 797 AAAGCAAAAGAACTAATCCAGGGTGAGATTAATTTAGTTTAAAAATTTAGTTAACAC 856
QY 601 GAGGAAAAA-GCTGCTGTGACAGCGAGTCACTGTTATCTTTACCTGCTGCTGCTGCTGCTGCT 659
Db 857 GAGGAAAAAAGGCTGCTGACAGCGAGTCACTGTTATCTTTACCTGCTGCTGCTGCTGCTGCT 916
QY 660 GTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 719
Db 917 GTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 976
QY 720 CCGCTATATAAATTCATATATTTTCCCTCCCGCTTTGAAATTTGCTGCTGCTGCTGCTGCTGCT 779
Db 977 CCGCTATATAAATTCATATATTTTCCCTCCCGCTTTGAAATTTGCTGCTGCTGCTGCTGCTGCT 1036
QY 780 CTTTCATCAGCGCTTTTGAATCTCCGCGGACTTCACAGAGAAAGAAAGAAAGAAAGAAAGAA 839
Db 1037 CTTTCATCAGCGCTTTTGAATCTCCGCGGACTTCACAGAGAAAGAAAGAAAGAAAGAAAGAA 1096
QY 840 AGAGAGAAAGTAAAGATTAATCCAGAGATTCATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 899
Db 1097 AGAGAGAAAGTAAAGATTAATCCAGAGATTCATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1156
QY 900 CATCT 959
Db 1157 CA---TCCT 1213
QY 960 TGTGATCTCGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1019
Db 1214 TGTGATCTCGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1273
QY 1020 TCTGTGAACTCCACTAATCTTTTCTTCCAGGTAATAGGACTTCTGATCTACTTTAT 1079
Db 1274 TCTGTGAACTCCACTAATCTTTTCTTCCAGGTAATAGGACTTCTGATCTACTTTAT 1333
QY 1080 TTAGCTCGATTAATAGCTTACCAGAAATTCGCTTGAATTCGATCTGATGAGATCCATGTTCTAGT 1139
Db 1334 TTAGCTCGATTAATAGCTTACCAGAAATTCGCTTGAATTCGATCTGATGAGATCCATGTTCTAGT 1393
QY 1140 TTACCTGGGAAATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1199
Db 1394 TTACCTGGGAAATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1453
QY 1200 TCTGACACTCTGCAATGTAGATTCGAATCTGGAACACTGTTTAAAGTTAGTGAAGTTGTTGTTG 1259

Db 1454 TCTGAACACTGTCAATGTTAGATTGAATCTGAACACTGTTTAAAGTTAGTGAAGTTGTTG 1513
QY 1260 TATAGATCTTCGAAACCTTAGGATTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1319
Db 1514 TATAGATCTTCGAAACCTTAGGATTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1573
QY 1320 GATTCATCAATCAGGCTTTTATTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1379
Db 1574 GATTCATCAATCAGGCTTTTATTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1633
QY 1380 AATTCGGCTG 1388
Db 1634 GGATCCCGG 1642

RESULT 5
AX392814
LOCUS AX392814 11127 bp DNA linear PAT 23-MAR-2002
DEFINITION Sequence 17 from Patent WO0215701.
ACCESSION AX392814
VERSION AX392814.1 GI:19700916
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Carotzi,N.B., Rabe,S.M., Miles,P.J., Warren,G.W. and de Haan,P.T.
TITLE Novel insecticidal toxins derived from Bacillus thuringiensis
JOURNAL insecticidal crystal proteins
Patent: WO 0215701-A 17 28-FEB-2002;
Syngenta Participations AG (CH)
FEATURES
Location/Qualifiers
1..11127
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="pZU578"
56..1475
/note="Actin 2 promoter U41998"
1485..1491
/note="synthetic nucleotide sequence encoding the toxin
portion of H04 plus the first 40 amino acids of the CryIAb
tail"
3859..5030
/note="SMAS promoter"
5052..6271
/note="PMI"

ORIGIN
Query Match 90.5%; Score 1306.2; DB 6; Length 11127;
Best Local Similarity 98.8%; Pred.No. 1.2e-273;
Matches 1368; Conservative 0; Mismatches 10; Indels 7; Gaps 5;
QY 1 ATTATGATCTCAATACATATGATACATATCTCATCTAGATCTAGGTTATCATATTATTAAG 60
Db 93 ATTATGATCTCAATACATATGATACATATCTCATCTAGATCTAGGTTATCATATTATTAAG 152
QY 61 AAAGTTTTCAGCAATATGNNACGACAAAATGGCTACACTCGATGTAATTTGGTATCTCAAC 120
Db 153 AAAGTTTTCAGCAATATGNNACGACAAAATGGCTAGACTCGATGTAATTTGGTATCTCAAC 212
QY 121 TCACATATTACTTATACCAACATTTAGTTAG-CAAAATTTAAACAATA-TTTTATGT 178
Db 213 TCACATATTACTTATACCAACATTTAGTTAGCAAAAATTTAAACAATA-TTTTATGT 272
QY 179 ATGCAAGAGTCAGCATATGTAATTTGATTCAGAAATCGTTTTCAGAGTTTCGGATCTAGT 238
Db 273 ATGCAAGAGTCAGCATATGTAATTTGATTCAGAAATCGTTTTCAGAGTTTCGGATCTAGT 332
QY 239 AGTAGCCATTTTAAATGTACATACATAATCGTGAATAGTG-ATATGATGAACATTTGAT 297
Db 333 AGTAGCCATTTTAAATGTACATACATAATCGTGAATAGTGATGATGATGATGATGATGATGAT 392

QY 298 CTTATTGTATATAATCCATAAACAACATCATGAAGACACTTTCTTTACGGCTCTGAATT 357
Db |||||
QY 393 CTTATTGTATATAATCCATAAACAACATCATGAAGACACTTTCTTTACGGCTCTGAATT 452
Db |||||
QY 358 AATTATGATACAAATCTTAATPAGAAACGAATTAATAGCTTGAATGTATGAATCTAA 417
Db |||||
QY 453 AATTATGACAAATCTTAATPAGAAACGAATTAATAGCTTGAATGTATGAATCTAA 512
Db |||||
QY 418 TTGAACAAGCCCAACGACGAGGACTTAACGTTGCCTGGAATGACTCGGTTTAAGTTAA 477
Db |||||
QY 513 TTGAACAAGCCCAACGACGAGGACTTAACGTTGCCTGGAATGACTCGGTTTAAGTTAA 572
Db |||||
QY 478 CACTAAAAAAGCGAGCTGTCTATGTAACACGCGATCGAGCAGGTACAGTCATGAAGCC 537
Db |||||
QY 573 CACTAAAAAAGCGAGCTGTCTATGTAACACGCGATCGAGCAGGTACAGTCATGAAGCC 632
Db |||||
QY 538 ATCAAGCAAAAGAACTAATCAAGGGGTGAGATGAATTAATAGTTTAAAAATTAAGTTAA 597
Db |||||
QY 633 ATCAAGCAAAAGAACTAATCAAGGGGTGAGATGAATTAATAGTTTAAAAATTAAGTTAA 692
Db |||||
QY 598 CACGAGGAAAA-GCTGTCTGACAGCCAGGTACAGTTATCTTACCTGTGGTGAATGA 656
Db |||||
QY 693 CACGAGGAAAAAGGCTGTCTGACAGCCAGGTACAGTTATCTTACCTGTGGTGAATGA 752
Db |||||
QY 657 TTCTGTCTGTCTGATTTAAATTAATTTTAAAAAGCCGAAATAAAAGTTGTAAGAGATA 716
Db |||||
QY 753 TTCTGTCTGTCTGATTTAAATTAATTTTAAAAAGCCGAAATAAAAGTTGTAAGAGATA 812
Db |||||
QY 717 AACCGGCTATATAAATCAATATATTTCTCCCGCTTTGAATGTCTGTCTCTCC 776
Db |||||
QY 813 AACCGGCTATATAAATCAATATATTTCTCCCGCTTTGAATGTCTGTCTCTCC 872
Db |||||
QY 777 TCACCTTTCATCAGCGTTTGAATCTCCGCGACTGTGACAGAGAAACAAGAGAGA 836
Db |||||
QY 873 TCACCTTTCATCAGCGTTTGAATCTCCGCGACTGTGACAGAGAAACAAGAGAGA 932
Db |||||
QY 837 CTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTTCTCCGTTTGAATCTCTCAA 896
Db |||||
QY 933 CTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTTCTCCGTTTGAATCTCTCAA 992
Db |||||
QY 897 TCTCATCTTCTCTCCGCTCTTTCTTCTCCAGGTAATPAGAACTTTCTGGACTACTTT 956
Db |||||
QY 993 TCTCA---TCTTCTCCGCTCTTTCTTCTCCAGGTAATPAGAACTTTCTGGACTACTTT 1049
Db |||||
QY 957 ATTTGCTGGATCTCGATCTTGTCTTCTCAATTTCCGTTGAGATCTGGAATCTGTTAATTT 1016
Db |||||
QY 1050 ATTTGCTGGATCTCGATCTTGTCTTCTCAATTTCCGTTGAGATCTGGAATCTGTTAATTT 1109
Db |||||
QY 1017 GGATCTGTGAACCTCCACTAAATCTTTTGGTTTAACTAGATCGATCTAAGTTGACCGAT 1076
Db |||||
QY 1110 GGATCTGTGAACCTCCACTAAATCTTTTGGTTTAACTAGATCGATCTAAGTTGACCGAT 1169
Db |||||
QY 1077 CAGTTAGCTCGATTAAGTACAGCAATTTGGCTTGACCTGTGAGAGATCCATGTTTC 1136
Db |||||
QY 1170 CAGTTAGCTCGATTAAGTACAGCAATTTGGCTTGACCTGTGAGAGATCCATGTTTC 1229
Db |||||
QY 1137 ATGTTAGCTGGGAATGATTTGTATATGTAATCTGAATCTGAACCTGTTTAACTAGATTT 1196
Db |||||
QY 1230 ATGTTAGCTGGGAATGATTTGTATATGTAATCTGAATCTGAACCTGTTTAACTAGATTT 1289
Db |||||
QY 1197 GAATCTGAACCTGTCAATGTTAGATTTGAATCTGAACCTGTTTAACTAGATTTAGATTT 1256
Db |||||
QY 1290 GAATCTGAACCTGTCAATGTTAGATTTGAATCTGAACCTGTTTAACTAGATTTAGATTT 1349
Db |||||
QY 1257 GTGATAGATTTCTTGAACCTTAGATTTGATGTTGATGTTGATGTTGAACAGATTT 1316
Db |||||
QY 1350 GTGATAGATTTCTTGAACCTTAGATTTGATGTTGATGTTGATGTTGAACAGATTT 1409
Db |||||
QY 1317 TCTGATTCATCAGCGTTTATTTGACTGTATTTGAACCTTTTGTGTTGCTGCTGAT 1376
Db |||||
QY 1410 TCTGATTCATCAGCGTTTATTTGACTGTATTTGAACCTTTTGTGTTGCTGCTGAT 1469
Db |||||

QY 1377 AAAAA 1381
Db |||||
QY 1470 AAAAA 1474
Db |||||
RESULT 6
AX463620
LOCUS AX463620
DEFINITION Sequence 17 from Patent WO0248335.
ACCESSION AX463620
VERSION AX463620.1 GI:21886380
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Meagher,R.B. and Li,Y.
TITLE Metal resistant plants and phytoremediation of environmental
JOURNAL contamination
Patent: WO 0248335-A 17 20-JUN-2002;
UNIVERSITY OF GEORGIA RESEARCH FOUNDATION, INC. (US)
FEATURES
Location/Qualifiers
1..3450
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Plasmid pEGSACT2"
ORIGIN
Query Match 89.9%; Score 1298.4; DB 6; Length 3450;
Best Local Similarity 98.1%; Pred. No. 6.7e-272;
Matches 1376; Conservative 0; Mismatches 18; Indels 8; Gaps 6;
QY 1 ATTATGATCTCAATACATGATACATATCTCATCTAGATCTAGGTATCATTTATCTATGTAAG 60
Db |||||
QY 9 ATGCTGATCTCAATACATGATACATATCTCATCTAGATCTAGGTATCATTTATGTAAG 68
Db |||||
QY 61 AAAGTTTTCGAGATATGNNAGCAAAATGGCTACACTCGATGTAATGGTATCTCAAC 120
Db |||||
QY 69 AAAGTTTTCGAGATATGNNAGCAAAATGGCTAGCTCGATGTAATGGTATCTCAAC 128
Db |||||
QY 121 TCACATTAATTAATACCAACATTAAGTAG-CAAAAATTTAAACAACATA-TTTTATGT 178
Db |||||
QY 129 TCACATTAATTAATACCAACATTAAGTAGTAGCAAAATTTAAACAACATA-TTTTATGT 188
Db |||||
QY 179 ATGCAAGAGTCAGCATATGTAATTAATTCAGATCGTTTTCAGAGTTCCGATGTAGT 238
Db |||||
QY 189 ATGCAAGAGTCAGCATATGTAATTAATTCAGATCGTTTTCAGAGTTCCGATGTAGT 248
Db |||||
QY 239 AGTAGCCATTTTAAATGTAATTAATTCAGATCGTTTTCAGAGTTCCGATGTAGT 297
Db |||||
QY 249 AGTAGCCATTTTAAATGTAATTAATTCAGATCGTTTTCAGAGTTCCGATGTAGT 308
Db |||||
QY 298 CTTATTGTATAAATATCCATAAACAACATCATGAAGACACTTTCTTTACGGCTCTGAATT 357
Db |||||
QY 309 CTTATTGTATAAATATCCATAAACAACATCATGAAGACACTTTCTTTACGGCTCTGAATT 368
Db |||||
QY 358 AATTATGATACAAATCTTAATAGAAAAAGAAATTAATAGTTGTAATGTATGAATCTAA 417
Db |||||
QY 369 AATTATGATACAAATCTTAATAGAAAAAGAAATTAATAGTTGTAATGTATGAATCTAA 428
Db |||||
QY 418 TTGAACAAGCCCAACGACGAGGACTTAACGTTGCCTGGAATGACTCGGTTTAAAGTTAA 477
Db |||||
QY 429 TTGAACAAGCCCAACGACGAGGACTTAACGTTGCCTGGAATGACTCGGTTTAAAGTTAA 488
Db |||||
QY 478 CACTAAAAAAGCGAGCTGTCTATGTAACACGCGATCGAGCAGGTACAGTCATGAAGCC 537
Db |||||
QY 489 CACTAAAAAAGCGAGCTGTCTATGTAACACGCGATCGAGCAGGTACAGTCATGAAGCC 548
Db |||||
QY 538 ATCAAGCAAAAGAACTAATCAAGGGGTGAGATGAATTAATAGTTTAAAAATTAAGTTAA 597
Db |||||
QY 549 ATCAAGCAAAAGAACTAATCAAGGGGTGAGATGAATTAATAGTTTAAAAATTAAGTTAA 608
Db |||||

QY 598 CACGAGGAAA-GCTGTCTGACAGCCAGGTCAAGTATCTTTACCTGTGTGTCGAATGA 656
 Db 609 CACGAGGAAAAGGCTGTCTGACAGCCAGGTCAAGTATCTTTACCTGTGTGTCGAATGA 668
 QY 657 TTCGTGTCTGTGATTTAAATTAATTTTGAAGGCGGAAATAAGTCTGAAGAGATA 716
 Db 669 TTCGTGTCTGTGATTTAAATTAATTTTGAAGGCGGAAATAAGTCTGAAGAGATA 728
 QY 717 AACCCGCTATATAAATTCATATATTTCTCCCGCTTTGAAATGTCTGTCTCTCC 776
 Db 729 AACCCGCTATATAAATTCATATATTTCTCCCGCTTTGAAATGTCTGTCTCTCC 788
 QY 777 TCACCTTCATCAGCGCTTTGAAATCTCCGCGGCTTGCACAGAGAAAGCAAGAGAGA 836
 Db 789 TCACCTTCATCAGCGCTTTGAAATCTCCGCGGCTTGCACAGAGAAAGCAAGAGAGA 848
 QY 837 CTAAGAGAAAGTAAAGAGATAATCCAGGAGATTCATTTCTCCGCTTTGAAATCTCTCAA 896
 Db 849 CTAAGAGAAAGTAAAGAGATAATCCAGGAGATTCATTTCTCCGCTTTGAAATCTCTCAA 908
 QY 897 TCTCANTCTTCTTCTCCGCTTTCTTCCAGAGTATAGAACTTCTGGAATCTACTTT 956
 Db 909 TCTCA---TCTTCTTCCGCTTTCTTCTCCAGGTAATAGAACTTCTGGAATCTACTTT 965
 QY 957 ATTGTGTCATCTCATCTTGTCTTCTCAATTTCTTGGATCTGGAATCTGTTAAATTT 1016
 Db 966 ATTGTGTCATCTCATCTTGTCTTCTCAATTTCTTGGATCTGGAATCTGTTAAATTT 1025
 QY 1017 GGATCTGTGAACTCTCACTAAATCTTTTGGTCTTCTAGATCTGATCTGATCTGATCT 1076
 Db 1026 GGATCTGTGAACTCTCACTAAATCTTTTGGTCTTCTAGATCTGATCTGATCTGATCT 1085
 QY 1077 CAGTGTAGCTGATATAGCTACAGAAATTTGGCTTGGATCTGATCTGATCTGATCTGATCT 1136
 Db 1086 CAGTGTAGCTGATATAGCTACAGAAATTTGGCTTGGATCTGATCTGATCTGATCTGATCT 1145
 QY 1137 ATGTACCTGGAAATGATTTGTATGTGAATGAAATCTGAACTGTGGAATGTAGATT 1196
 Db 1146 ATGTACCTGGAAATGATTTGTATGTGAATGAAATCTGAACTGTGGAATGTAGATT 1205
 QY 1197 GAATCTGAACCTGTCAATGTAGATGAATCTGAACTGTGGAATGTAGATT 1255
 Db 1206 GAATCTGAACCTGTCAATGTAGATGAATCTGAACTGTGGAATGTAGATT 1265
 QY 1256 TGTGTATAGATTTTGAACCTGTAGATTTGTAGTGTGATCTGAACTGTGGAATGTAGATT 1315
 Db 1266 TGTGTATAGATTTTGAACCTGTAGATTTGTAGTGTGATCTGAACTGTGGAATGTAGATT 1325
 QY 1316 TTCGATTCATTCAGGCTTTATTTGATCTGATGAACTCTTTTGTGTGTTGCACTCA 1375
 Db 1326 TTCGATTCATTCAGGCTTTATTTGATCTGATGAACTCTTTTGTGTGTTGCACTCA 1385
 QY 1376 TAAATAATGGCTGAGGCTGACG 1397
 Db 1386 TAAACCATGGCTGAGGCTGACG 1407

RESULT 7
 AX463619
 LOCUS AX463619 3408 bp DNA linear PAT 15-JUL-2002
 DEFINITION Sequence 16 from Patent WO248335.
 ACCESSION AX463619
 VERSION AX463619.1 GI:21886379
 KEYWORDS

SOURCE synthetic construct
 ORGANISM synthetic construct
 1 artificial sequences.
 REFERENCE Meagher,R.B. and Li,Y.
 AUTHORS Metal resistant plants and phytoremediation of environmental
 TITLE contamination
 JOURNAL Patent: WO 0248335-A 16 20-JUN-2002;
 UNIVERSITY OF GEORGIA RESEARCH FOUNDATION, INC.. (US)

FEATURES
 source
 1. 3408
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xrefs="taxon:32630"
 /note="Plasmid pATPCSACT2"
 ORIGIN
 Query Match 89.8%; Score 1297.4; DB 6; Length 3408;
 Best Local Similarity 98.5%; Pred. No. 1.1e-271;
 Matches 1372; Conservative 0; Mismatches 13; Indels 8; Gaps 6;
 QY 1 AATTGATCTCAAAATACATTTGATACATATCTCATCTAGATCTAGGTATCATATATGTAAG 60
 Db 9 ATGCTGATCTCAAAATACATTTGATACATATCTCATCTAGATCTAGGTATCATATATGTAAG 68
 QY 61 AAGGTTTTCAGCAATATGNNACGACCAAAATGGCTACCTCGATGTAATTTGGTATCTCAAC 120
 Db 69 AAGGTTTTCAGCAATATGNNACGACCAAAATGGCTAGCTCGATGTAATTTGGTATCTCAAC 128
 QY 121 TCAACATTAATCTATACCAAAACATTTAGTTAG-CAAAAATTTAAACAACTA-TTTTTATGT 178
 Db 129 TCAACATTAATCTATACCAAAACATTTAGTTAGCACAAAATTTAAACAACTA-TTTTTATGT 188
 QY 179 ATGCAAGAGTCAGCATATGTAATTAATTTGATTTGAAATCGTTTTGACGAGTTCGGATGAGT 238
 Db 189 ATGCAAGAGTCAGCATATGTAATTAATTTGATTTGAAATCGTTTTGACGAGTTCGGATGAGT 248
 QY 239 AGTAGCCATTAATTAATGTACATCTTAATCGTAATAGTG-ATATGATGAAACATTTGAT 297
 Db 249 AGTAGCCATTAATTAATGTACATCTTAATCGTAATAGTGAAATGATGAAACATTTGAT 308
 QY 298 CTTATTTGATAAATATCCATAAACAATCATGAAGAACA-TTTTTCACGGGTCTGAAT 357
 Db 309 CTTATTTGATAAATATCCATAAACAATCATGAAGAACA-TTTTTCACGGGTCTGAAT 368
 QY 358 AATTATGATACATTTCAATAGAAAGCAATTAATTTACGTTGAATTTGATGAAATCTAA 417
 Db 369 AATTATGATACATTTCAATAGAAAGCAATTAATTTACGTTGAATTTGATGAAATCTAA 428
 QY 418 TTGAACAAGCCAAACCAACGACGAGGACTTAACGTTGCTGGATTTGACTCGGTTTAAAGTTAAC 477
 Db 429 TTGAACAAGCCAAACCAACGACGAGGACTTAACGTTGCTGGATTTGACTCGGTTTAAAGTTAAC 488
 QY 478 CACTAAAAAAGCGAGCTGTCTGTAAACAACGCGGATCGACAGGTCAACAGTCATGAAGCC 537
 Db 489 CACTAAAAAAGCGAGCTGTCTGTAAACAACGCGGATCGACAGGTCAACAGTCATGAAGCC 548
 QY 538 ATCAAGCAAAAGAACTAATCCAAAGGGGTGAGATGATTAATTTAGTTTAAATTTAGTTAA 597
 Db 549 ATCAAGCAAAAGAACTAATCCAAAGGGGTGAGATGATTAATTTAGTTTAAATTTAGTTAA 608
 QY 598 CACGAGGAAAA-GCTGTCTGACAGCCAGGTCAAGTATCTTTTACCTGTGTGTCGAATGA 656
 Db 609 CACGAGGAAAAAGGCTGTCTGACAGCCAGGTCAAGTATCTTTTACCTGTGTGTCGAATGA 668
 QY 657 TTGCTGTCTGTGATTTTAATTAATTTTGAAGGCGGAAATAAGTCTGAAGAGATA 716
 Db 669 TTGCTGTCTGTGATTTTAATTAATTTTGAAGGCGGAAATAAGTCTGAAGAGATA 728
 QY 717 AACCCGCTATATAAATTCATATATTTTCTCCCGCTTTGAAATGTCTGTGTCTCTCC 776
 Db 729 AACCCGCTATATAAATTCATATATTTTCTCCCGCTTTGAAATGTCTGTGTCTCTCC 788
 QY 777 TCACCTTCATCAGCGCTTTGAAATCTCCGCGGCTTGCACAGAGAAAGCAAGAGAGA 836
 Db 789 TCACCTTCATCAGCGCTTTGAAATCTCCGCGGCTTGCACAGAGAAAGCAAGAGAGA 848
 QY 837 CTAAGAGAAAGTAAAGAGATAATCCAGGAGATTCATTTCTCCGCTTTGAAATCTCTCAA 896
 Db 849 CTAAGAGAAAGTAAAGAGATAATCCAGGAGATTCATTTCTCCGCTTTGAAATCTCTCAA 908
 QY 897 TCTCATCTTCTTCTTCTCCGCTTTCTTCTTCCAGGTAATAGAACTTTCTGGAATCTACTTT 956

Db 1206 GAATCTGAACACTGCTCAATGTTAGATTGAATCTGAACACTGTTTAAAGTTAGATGAAGTT 1265
 QY 1256 TGTGTATAGATCTTCGAAACCTTAGGATTTGTAGTGTGTCGTAAGTGAACAGAAAGCTAT 1315
 Db 1266 TGTGTATAGATCTTCGAAACCTTAGGATTTGTAGTGTGTCGTAAGTGAACAGAAAGCTAT 1325
 QY 1316 TTCTGATTCATCAGGGTTTATTCAGCTGATTAAGTGAATTTTGTGTTTGTGCTGATCA 1375
 Db 1326 TTCTGATTCATCAGGGTTTATTCAGCTGATTAAGTGAATTTTGTGTTTGTGCTGATCA 1385
 QY 1376 TAAAAATGGC 1386
 Db 1386 TAAACATGGC 1396

RESULT 9
 ATU41998
 LOCUS 3172 bp DNA linear PLN 30-NOV-1996
 DEFINITION Arabidopsis thaliana actin 2 (ACT2) gene, complete cds.
 ACCESSION U41998
 VERSION U41998.1 GI:1669386
 KEYWORDS
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE 1 (bases 1 to 3172)
 An, Y.-Q., McDowell, J. M., Huang, S., McKinney, E. C., Chambliss, S. and
 Meagher, R. B.
 TITLE Strong, constitutive expression of the Arabidopsis ACT2/ACT8 actin
 subclass in vegetative tissues
 JOURNAL Plant J. 10 (1), 107-121 (1996)
 MEDLINE 96314863
 PUBMED 8759981
 REFERENCE 2 (bases 1 to 3172)
 An, Y.-Q.
 TITLE Direct Submission
 JOURNAL Submitted (05-DEC-1995) Yong-Qiang An, Department of Genetics,
 University of Georgia, Athens, GA 30602, USA
 FEATURES
 source
 1. 3172
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /db_xref="taxon:3702"
 gene 563..2849
 /gene="ACT2"
 563..568
 /gene="ACT2"
 TATA_signal join(597..764,1207..1670,1749..2362,2449..2849)
 mRNA /product="actin 2"
 /protein_id="AA37098.1"
 /db_xref="GI:1669387"
 CDS join(1217..1670,1749..2362,2449..2514)
 /gene="ACT2"
 /codon_start=1
 /product="actin 2"
 /protein_id="AA37098.1"
 /db_xref="GI:1669387"

WVGWNNQKJYVGDQKQKGLILKPIEHGVVSNWDDMEKILWHITFYNELRIAPPE
 HPVLITAPLNPRANKRMQIMFETNSPAMYAIQVLSYVAGRTTGVLDSDGG
 VSHVTPYEFSLPHAILRLDLAAGLDLTKMLITERTGYMTFTAEIREIVRDKKL
 SFVAVDYEMETSKTSSEIKNYELPDQVITIGAFRCPEVLPQSPFVMEAGI
 HETTVNSIMKCDVIDRKDLVGNIVLSGGTTFPSGIDRMSKETALAPSSMKIKVAP
 PERKYSVWIGGSLASLSTFQOMWISKAYDEAGPIVHRKF"

ORIGIN

Query Match 85.5%; Score 1234.2; DB 8; Length 3172;
 Best Local Similarity 99.1%; Pred. No. 6.1e-258;
 Matches 1263; Conservative 0; Mismatches 8; Indels 4; Gaps 2;
 QY 164 CAACTATTTTATGATGCAAGCTCAGCATATGATATGATTCAGATCGTTTTCAC 223
 Db 1 CAACTATTTTATGATGCAAGCTCAGCATATGATATGATTCAGATCGTTTTCAC 60

QY 224 GAGTTCGAGTGTAGTAGCCCAATTATTTAATGTACATCTAATCGTGAATAGTGATG 283
 Db 61 GAGTTCGAGTGTAGTAGCCCAATTATTTAATGTACATCTAATCGTGAATAGTGATG 120
 QY 284 ATGAACATCTGATCTTATTTGTATTAATCCATAACACATCATGAAGACACTTCTT 343
 Db 121 ATGAACATCTGATCTTATTTGTATTAATCCATAACACATCATGAAGACACTTCTT 180
 QY 344 TCAGGCTCGAATTAATTAATGATACAATTCCTAATAGAAAACGAATTAATTAAGTTGAAT 403
 Db 181 TCAGGCTCGAATTAATTAATGATACAATTCCTAATAGAAAACGAATTAATTAAGTTGAAT 240
 QY 404 TGTATGAATCTAATTTGAACAGCCCAACACGACGAGGACTAAAGTTCGCTGGAATGACT 463
 Db 241 TGTATGAATCTAATTTGAACAGCCCAACACGACGAGGACTAAAGTTCGCTGGAATGACT 300
 QY 464 CGGTTTAAAGTTAACCTTAAAAAAACGAGCTGTCAACACGCGATCGAGCAGGTC 523
 Db 301 CGGTTTAAAGTTAACCTTAAAAAAACGAGCTGTCAACACGCGATCGAGCAGGTC 360
 QY 524 ACAGTCATGAAGCCCATCAAGCAAAAGAACTAATCCAAGGGCTGAGATGATTAATAGTT 583
 Db 361 ACAGTCATGAAGCCCATCAAGCAAAAGAACTAATCCAAGGGCTGAGATGATTAATAGTT 420
 QY 584 TAAANNTAGTTAACAGAGGAGGAAA-CTCTCTGACAGCGAGTCACTATCTTTACC 642
 Db 421 TAAANNTAGTTAACAGAGGAGGAGGCTGTCTGACAGCGAGTCACTATCTTTACC 480
 QY 643 TGTGGTCCGAATGATTCGTCTGTCTGATTAATTAATTAATTTTGAAGGCCGAAATAA 702
 Db 481 TGTGGTCCGAATGATTCGTCTGTCTGATTAATTAATTAATTTTGAAGGCCGAAATAA 540
 QY 703 AGTTGTAGAGATTAACCCGCTATATAAATTCATATTTCTCCGCCGTTTGAATTG 762
 Db 541 AGTTGTAGAGATTAACCCGCTATATAAATTCATATTTCTCCGCCGTTTGAATTG 600
 QY 763 TCTCGTGTCTCTCTCACTTCTCATCAGCCGTTTGAATCTCCGCGACTTGACAGAGAG 822
 Db 601 TCTCGTGTCTCTCTCACTTCTCATCAGCCGTTTGAATCTCCGCGACTTGACAGAGAG 660
 QY 823 AACAGAGAAGAACTAAGAGAGAAGTAAGAGATAATCCAGGAGATTCATCTCCGTTT 882
 Db 661 AACAGAGAAGAACTAAGAGAGAAGTAAGAGATAATCCAGGAGATTCATCTCCGTTT 720
 QY 883 TGAATCTCTCAATCTCATCTCTTCTTCGCTCTTCTTCCAGAGTAATAGGACTT 942
 Db 721 TGAATCTCTCAATCTCA--TCTTCTCCGCTCTTCTTCCAGAGTAATAGGAACTT 777
 QY 943 TCTGGATCTACTTTATTTGCTGGATCTCGATCTTGTCTTCTCAATTTCTTGAGATCTGG 1002
 Db 778 TCTGGATCTACTTTATTTGCTGGATCTCGATCTTGTCTTCTCAATTTCTTGAGATCTGG 837
 QY 1003 AATTGCTTAAATTTGGATCTGTGAACCTCCAATAAATCTTTGGTTTACTAGATCGAT 1062
 Db 838 AATTGCTTAAATTTGGATCTGTGAACCTCCAATAAATCTTTGGTTTACTAGATCGAT 897
 QY 1063 CTAAGTTGACCGATCAGTGTAGCTCCATTTATAGTACAGAAATTTGGCTTACGCTTGATGG 1122
 Db 898 CTAAGTTGACCGATCAGTGTAGCTCCATTTATAGTACAGAAATTTGGCTTACGCTTGATGG 957
 QY 1123 AGAGATCCAATGTCATGTTTACCTGGAAAATGATTTGTATGTGAATTTGAATCTGAACT 1182
 Db 958 AGAGATCCAATGTCATGTTTACCTGGAAAATGATTTGTATGTGAATTTGAATCTGAACT 1017
 QY 1183 GTTGAAGTTAGATGAATCTGAACACTGTGCAATGTTAGATGAATCTGAACACTGTTAA 1242
 Db 1018 GTTGAAGTTAGATGAATCTGAACACTGTGCAATGTTAGATGAATCTGAACACTGTTAA 1077
 QY 1243 GTTAGATGAAGTTTGTGTATAGATTTCTTCCGAACTTAGGATTTGTAGTCTGCTGCTTG 1302
 Db 1078 GTTAGATGAAGTTTGTGTATAGATTTCTTCCGAACTTAGGATTTGTAGTCTGCTGCTTG 1137

QY 1303 AACAGAACATATCTGATTCATACAGGGTTTATTTGACTGTATTAATGAATCTTTTGG 1362
| | | | |
Db 1138 AACAGAACATATCTGATTCATACAGGGTTTATTTGACTGTATTAATGAATCTTTTGG 1197
| | | | |
QY 1363 TGTTTCAGCTCATAAAAATGCTGAGCTGACGATTAATCAACCAATCGTGTGACAA 1422
| | | | |
Db 1198 TGTTTCAGCTCATAAAAATGCTGAGCTGACGATTAATCAACCAATCGTGTGACAA 1257
| | | | |
QY 1423 TGGTACTGGAATGGT 1437
| | | | |
Db 1258 TGGTACCGGTATGGT 1272
| | | | |

RESULT 10
AX461227 1217 bp DNA linear PAT 08-JUL-2002
LOCUS Sequence 156 from Patent WO0198480.
DEFINITION AX461227
ACCESSION AX461227
VERSION AX461227.1 GI:21726435
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
1
REFERENCE Budworth, P., Brown, D., Chang, H.S., Zhu, T., Han, B., Wang, X. and
Cooper, B.
PROMOTERS for regulation of plant gene expression
TITLE Patent: WO 0198480-A 156 27-DEC-2001;
JOURNAL Syngenta Participations AG (CH)
FEATURES Location/Qualifiers
Source 1..1217
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

Query Match 82.0%; Score 1184; DB 6; Length 1217;
Best Local Similarity 99.3%; Pred. No. 5.6e-247;
Matches 1211; Conservative 0; Mismatches 5; Indels 4; Gaps 2;

QY 164 CAACATATTTTATGATGCAAGAGTCAGCATATGTATAATTAATGATCAGAAATCGTTTGAC 223
| | | | |
Db 1 CAACATATTTTATGATGCAAGAGTCAGCATATGTATAATTAATGATCAGAAATCGTTTGAC 60
| | | | |
QY 224 GAGTTCGGATGTAGTAGTACGCAATTAATTAATGATGATCTAATCTGTAATGATGATGATG 283
| | | | |
Db 61 GAGTTCGGATGTAGTAGTACGCAATTAATTAATGATGATCTAATCTGTAATGATGATGATG 120
| | | | |
QY 284 ATGAACATTTGATCTTATTTGTAATAATATCAATAACATCAATCAATCAATCAATCAAT 343
| | | | |
Db 121 ATGAACATTTGATCTTATTTGTAATAATATCAATAACATCAATCAATCAATCAATCAAT 180
| | | | |
QY 344 TCAGGGTCTGAATTAATTAATGATACAAATTTCTAATAGAAAAAGAAATTAATTAATGATGAT 403
| | | | |
Db 181 TCAGGGTCTGAATTAATTAATGATACAAATTTCTAATAGAAAAAGAAATTAATTAATGATGAT 240
| | | | |
QY 404 TGATGAATCTAATTAATGACAGCAACCAACGAGGACTAATCGTTCGCTGGATTTGACT 463
| | | | |
Db 241 TGATGAATCTAATTAATGACAGCAACCAACGAGGACTAATCGTTCGCTGGATTTGACT 300
| | | | |
QY 464 CGGTTTAAGTTAACCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 523
| | | | |
Db 301 CGGTTTAAGTTAACCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 360
| | | | |
QY 524 ACAGTCATGAGCCATCAAGCAAAAGAACTAATCAAGGGGTGAGATGATTAATTAATGAT 583
| | | | |
Db 361 ACAGTCATGAGCCATCAAGCAAAAGAACTAATCAAGGGGTGAGATGATTAATTAATGAT 420
| | | | |
QY 584 TAAAAATAGTTAAACAGAGGAAAAA-GCTGTCTGACAGCGAGTCAAGTTATCTTTTACC 642
| | | | |
Db 421 TAAAAATAGTTAAACAGAGGAAAAAGGCTGTCTGACAGCGAGTCAAGTTATCTTTTACC 480
| | | | |

QY 643 TGTGTCGAAATGATTCGTGCTGTGCGATTTTAATTAATTTTGAAGGCGGAAATATA 702
| | | | |
Db 481 TGTGTCGAAATGATTCGTGCTGTGCGATTTTAATTAATTTTGAAGGCGGAAATATA 540
| | | | |
QY 703 AGTTGTAAGAGATAAACCCGCTATATAAATTCATATATTTTCTCCCGCTTTGAATG 762
| | | | |
Db 541 AGTTGTAAGAGATAAACCCGCTATATAAATTCATATATTTTCTCCCGCTTTGAATG 600
| | | | |
QY 763 TCTCGTTGCTCTCTCACTTTTCATCAGCGGTTTGAATCTCCGCGCACTTGACAGAGAAG 822
| | | | |
Db 601 TCTCGTTGCTCTCTCACTTTTCATCAGCGGTTTGAATCTCCGCGCACTTGACAGAGAAG 660
| | | | |
QY 823 AACAGGAAGAGATTAAGAGAGAAAGTAAGAGATAATCCAGAGATTCATCTCCGTTT 882
| | | | |
Db 661 AACAGGAAGAGATTAAGAGAGAAAGTAAGAGATAATCCAGAGATTCATCTCCGTTT 720
| | | | |
QY 883 TGAATCTTCTCAATCTCATCTCTCTTCTTCCGCTCTTCTTCCAGAGTAATAGAACTT 942
| | | | |
Db 721 TGAATCTTCTCAATCTCA--TCTTCTTCCGCTCTTCTTCCAGAGTAATAGAACTT 777
| | | | |
QY 943 TCTCGATCTACTTTATTTGCTGAGATCTGATCTTGTGTTTCTCAATTTCTCGAGATCG 1002
| | | | |
Db 778 TCTCGATCTACTTTATTTGCTGAGATCTGATCTTGTGTTTCTCAATTTCTCGAGATCG 837
| | | | |
QY 1003 AATTGCTTTAATTTGATCTGTGAACCTCCACTAAATCTTTTCTGTTTACTAGAACTGAT 1062
| | | | |
Db 838 AATTGCTTTAATTTGATCTGTGAACCTCCACTAAATCTTTTCTGTTTACTAGAACTGAT 897
| | | | |
QY 1063 CTAAGTTGACCGATCAGTTAGCTCGATTTAGCTACAGAAATTTGGCTTGACCTTGATGG 1122
| | | | |
Db 898 CTAAGTTGACCGATCAGTTAGCTCGATTTAGCTACAGAAATTTGGCTTGACCTTGATGG 957
| | | | |
QY 1123 AGAGATCCATGTTCTGTTTACCTGGGAATGATTTGTATATGTAATGAAATCTGAAT 1182
| | | | |
Db 958 AGAGATCCATGTTCTGTTTACCTGGGAATGATTTGTATATGTAATGAAATCTGAAT 1017
| | | | |
QY 1183 GTTCAAGTTAGATTTGAATCTGAACCTGTAATGTAATGTAATGTAATGTAATGTAAT 1242
| | | | |
Db 1018 GTTCAAGTTAGATTTGAATCTGAACCTGTAATGTAATGTAATGTAATGTAATGTAAT 1077
| | | | |
QY 1243 GTTAGATGAAGTTGTTGTAATGATTTCTCGAAACCTTAGGATTTGTAGTGTGACGTTG 1302
| | | | |
Db 1078 GTTAGATGAAGTTGTTGTAATGATTTCTCGAAACCTTAGGATTTGTAGTGTGACGTTG 1137
| | | | |
QY 1303 AACAGAAAGCTATTTCTGATTCATCAGGGTTTATTTGACTGTATTAATGAATCTTTTGTG 1362
| | | | |
Db 1138 AACAGAAAGCTATTTCTGATTCATCAGGGTTTATTTGACTGTATTAATGAATCTTTTGTG 1197
| | | | |
QY 1363 TGTTTGCGAGCTCATAAAAA 1382
| | | | |
Db 1198 TGTTTGCGAGCTCATAAAAA 1217
| | | | |

RESULT 11
AR236191 1219 bp DNA linear PAT 20-DEC-2002
LOCUS Sequence 9 from patent US 6462258.
DEFINITION AR236191
ACCESSION AR236191
VERSION AR236191.1 GI:27280000
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1219)
AUTHORS Fincher, K.L. and Wilkinson, J.Q.
TITLE Plant expression constructs
JOURNAL Patent: US 6462258-A 9 08-OCT-2002;
FEATURES Location/Qualifiers
source 1..1219
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 81.4%; Score 1174.8; DB 6; Length 1219;
Best Local Similarity 98.7%; Pred. No. 5.6e-245;
Matches 1206; Conservative 0; Mismatches 12; Indels 4; Gaps 2;

QY 164 CAACTATTTTATGATGCAAGAGTCAGCATATGTATAATTTGATTCAGATCGTTTTCAC 223
DB 1 CAACTATTTTATGATGCAAGAGTCAGCATATGTATAATTTGATTCAGATCGTTTTCAC 60

QY 224 GAGTTCGGATGTAGTAGCCATTTTAAATGTACATACATTAATCGTGAATAGTATG 293
DB 61 GAGTTCGGATGTAGTAGCCATTTTAAATGTACATACATTAATCGTGAATAGTATG 120

QY 284 ATGAACAATCTATCTTATTTGATTAATATCCATAAACAACATCATGAAGACACTTTCTT 343
DB 121 ATGAACAATCTATCTTATTTGATTAATATCCATAAACAACATCATGAAGACACTTTCTT 180

QY 344 TCAGGCTCTGAATTAATATGATCAATTTCTAATAGAAAAGAAATTAATTCGTTGAAT 403
DB 181 TCAGGCTCTGAATTAATATGATCAATTTCTAATAGAAAAGAAATTAATTCGTTGAAT 240

QY 404 TGTATGAATCTAATTTGACAGCAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 463
DB 241 TGTATGAATCTAATTTGACAGCAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 300

QY 464 CGGTTTAAAGTTTAAACCACTAAATAAATGATGATGATGATGATGATGATGATGATGATG 523
DB 301 CGGTTTAAAGTTTAAACCACTAAATAAATGATGATGATGATGATGATGATGATGATGATG 360

QY 524 ACAGTCATGAAGCAATCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 582
DB 361 ACAGTCATGAAGCAATCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 420

QY 584 TAAATATGATTTACACAGGAGGAAA-GCTGTCTGACAGCAGGTCAGGTTATCTTTACC 642
DB 421 TAAATATGATTTACACAGGAGGAAAAGGCTGTCTGACAGCAGGTCAGGTTATCTTTACC 480

QY 643 TGTGTCGAAATGATGTCGTCTGTCGATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 702
DB 481 TGTGTCGAAATGATGTCGTCTGTCGATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 540

QY 703 AGTTGTAGAGATAAACCCTCTATATAATTAATTAATTAATTAATTAATTAATTAATTAAT 762
DB 541 AGTTGTAGAGATAAACCCTCTATATAATTAATTAATTAATTAATTAATTAATTAATTAAT 600

QY 763 TCTCGTGTCT 822
DB 601 TCTCGTGTCT 660

QY 823 AACAGGAGAGAGACTAAG 882
DB 661 AACAGGAGAGAGACTAAG 720

QY 883 TGAATCT 942
DB 721 TGAATCT 777

QY 943 TCTGGATCTACTTTATTTGTCGATCTCGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1002
DB 778 TCTGGATCTACTTTATTTGTCGATCTCGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 837

QY 1003 AATTCTGTTTAAATTTGATGTAACCTTCACTAATATCTTTTGGTTTATAGAAATCGAT 1062
DB 838 AATTCTGTTTAAATTTGATGTAACCTTCACTAATATCTTTTGGTTTATAGAAATCGAT 897

QY 1063 CTAAGTTCAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1122
DB 898 CTAAGTTCAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 957

QY 1123 AGAGATCCATGTTTCACTGACCTGGGAAATGATTTGATATGATGATGATGATGATGATGATG 1182
DB 958 AGAGATCCATGTTTCACTGACCTGGGAAATGATTTGATATGATGATGATGATGATGATGATG 1017

QY 1183 GTTGAAGTTAGATTTGAATCTGAACACTCTCAATGTTAGATTTGAATCTGAACACTGTTTAA 1242
DB 1018 GTTGAAGTTAGATTTGAATCTGAACACTCTCAATGTTAGATTTGAATCTGAACACTGTTTAA 1077

QY 1243 GTTGAATGAAGTTTGTGTATAGATTTCTTCGAAACCTTAGGATTTGTAGTGTCTAGCTTG 1302
DB 1078 GTTGAATGAAGTTTGTGTATAGATTTCTTCGAAACCTTAGGATTTGTAGTGTCTAGCTTG 1137

QY 1303 AACAGAAAGCTATTTCTGATTTCAATCAGGGTTTATTTGACTGTATTGAACTCTTTTGTG 1362
DB 1138 AACAGAAAGCTATTTCTGATTCATCAGGGTTTATTTGACTGTATTGAACTCTTTTGTG 1197

QY 1363 TGTTCGAGCTCATAAAAAATG 1384
DB 1198 TGTTCGAGCTCATCAATG 1219

RESULT 12
AR438153
LOCUS AR438153 1219 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 9 from patent US 6660911.
ACCESSION AR438153
VERSION AR438153.1 GI:40204655
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1219)
AUTHORS Fincher,K.L., Flasinaki,S. and Wilkinson,J.O.
TITLE Plant expression constructs
JOURNAL Patent: US 6660911-A 9 09-DEC-2003;
FEATURES
Location/Qualifiers
1..1219
source /organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 81.4%; Score 1174.8; DB 6; Length 1219;
Best Local Similarity 98.7%; Pred. No. 5.6e-245;
Matches 1206; Conservative 0; Mismatches 12; Indels 4; Gaps 2;

QY 164 CAACTATTTTATGATGCAAGAGTCAGCATATGTATAATTTGATTCAGATCGTTTTCAC 223
DB 1 CAACTATTTTATGATGCAAGAGTCAGCATATGTATAATTTGATTCAGATCGTTTTCAC 60

QY 224 GAGTTCGGATGTAGTAGCCATTTTAAATGTACATACATTAATCGTGAATAGTATG 283
DB 61 GAGTTCGGATGTAGTAGCCATTTTAAATGTACATACATTAATCGTGAATAGTATG 120

QY 284 ATGAACAATCTTATTTGATTAATATCCATAAACAACATCATGAAGACACTTTCTT 343
DB 121 ATGAACAATCTTATTTGATTAATATCCATAAACAACATCATGAAGACACTTTCTT 180

QY 344 TCAGGCTCTGAATTAATATGATCAATTTCTAATAGAAAAGAAATTAATTCGTTGAAT 403
DB 181 TCAGGCTCTGAATTAATATGATCAATTTCTAATAGAAAAGAAATTAATTCGTTGAAT 240

QY 404 TGTATGAATCTAATTTGACAGCAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 463
DB 241 TGTATGAATCTAATTTGACAGCAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 300

QY 464 CGGTTTAAAGTTTAAACCACTAAATAAATGATGATGATGATGATGATGATGATGATGATG 523
DB 301 CGGTTTAAAGTTTAAACCACTAAATAAATGATGATGATGATGATGATGATGATGATGATG 360

QY 524 ACAGTCATGAAGCAATCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 583
DB 361 ACAGTCATGAAGCAATCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 420

QY 584 TAAATATGATTTACACAGGAGGAAA-GCTGTCTGACAGCAGGTCAGGTTATCTTTACC 642
DB 421 TAAATATGATTTACACAGGAGGAAAAGGCTGTCTGACAGCAGGTCAGGTTATCTTTACC 480

QY 1123 AGAGATCATGTTTACCTGGGAAATGATTGTTATATGTGAATGAATCTGAAT 1182
DB 958 AGAGATCATGTTTACCTGGGAAATGATTGTTATATGTGAATGAATCTGAAT 1017
QY 1183 GTTGAAGTTAGATGAATCTGAACACTGTCAATGTTTAGATTGAATCTGAACACTGTTAA 1242
DB 1018 GTTGAAGTTAGATGAATCTGAACACTGTCAATGTTTAGATTGAATCTGAACACTGTTAA 1077
QY 1243 GTTAGATGAAGTTGTTGTTATAGATTCTTCCGAAACCTTAGATTGTTAGTGTGTACGTTG 1302
DB 1078 GTTAGATGAAGTTGTTGTTATAGATTCTTCCGAAACCTTAGATTGTTAGTGTGTACGTTG 1137
QY 1303 AACAGAAAGCTATTTCTGATCAATCAGGGTTTATTTAGTGTATTCGAATCTTTTCTG 1362
DB 1138 AACAGAAAGCTATTTCTGATCAATCAGGGTTTATTTAGTGTATTCGAATCTTTTCTG 1197
QY 1363 TGTTTGCAGCTCATAAAAATG 1384
DB 1198 TGTTTGCAGCACTCACCATG 1219
RESULT 14
AR236212
LOCUS 1742 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 30 from patent US 6462258.
ACCESSION AR236212
VERSION AR236212.1 GI:27280021
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1742)
AUTHORS Finch, K.L. and Wilkinson, J.Q.
TITLE Plant expression constructs
JOURNAL Patent: US 6462258-A 30 08-OCT-2002;
FEATURES
source Location/Qualifiers
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 81.3%; Score 1174; DB 6; Length 1742;
Best Local Similarity 99.3%; Pred. No. 8e-245;
Matches 1201; Conservative 0; Mismatches 5; Indels 4; Gaps 2;
QY 164 CAACATATTTATGTATGCAAGAGTCAGCATATGTATTAATTCATTCAGATCGTTTGAC 223
DB 533 CAACATATTTATGTATGCAAGAGTCAGCATATGTATTAATTCATTCAGATCGTTTGAC 592
QY 224 GAGTTCGGATGTAGTAGCCATTATTTAATGTACATATAATCGTGAATAGTGATG 283
DB 593 GAGTTCGGATGTAGTAGCCATTATTTAATGTACATATAATCGTGAATAGTGATG 652
QY 284 ATGAACATGTATCTTATGTATTAATATCCATAAACAATCATGAAAGACACTTCTT 343
DB 653 ATGAACATGTATCTTATGTATTAATATCCATAAACAATCATGAAAGACACTTCTT 712
QY 344 TCAGGCTGTAATTAATGTATTAATCTTAATAGAAACCAATTAATTCAGTTGAAT 403
DB 713 TCAGGCTGTAATTAATGTATTAATCTTAATAGAAACCAATTAATTCAGTTGAAT 772
QY 404 TGATGAAATCTTAATTTGAACAAGCCAAACAGCAGGAGCTAACCTGCTCGATTGACT 463
DB 773 TGATGAAATCTTAATTTGAACAAGCCAAACAGCAGGAGCTAACCTGCTCGATTGACT 832
QY 464 CGGTTAAGTTAACTCAATAAAAAAAGAGCTGTATGTAACACCGGATCGAGAGGTC 523
DB 833 CGGTTAAGTTAACTCAATAAAAAAAGAGCTGTATGTAACACCGGATCGAGAGGTC 892
QY 524 ACAGTCACTCAAGCCATCAAGCAAGAAAGACTAATCCAGGGGTGAGATGATTAATAGTT 583
DB 893 ACAGTCACTCAAGCCATCAAGCAAGAAAGACTAATCCAGGGGTGAGATGATTAATAGTT 952

QY 584 TAAATATAGTTAAACAGGGAAGA--GCTGTCTGACAGCAGGTCAGCTTATCTTTACC 642
DB 953 TAAATATAGTTAAACAGGGAAGAAGCTGTCTGACAGCAGGTCAGCTTATCTTTACC 1012
QY 643 TGTGTCGAAATGATTCGTGTCGTGATTTAATTTATTTTGAAGCCGAAATAA 702
DB 1013 TGTGTCGAAATGATTCGTGTCGTGATTTAATTTATTTTGAAGCCGAAATAA 1072
QY 703 AGTTGTAGAGATAAACCAGCCCTATATAAATCATATATATTTCTCCCGCTTTGAATG 762
DB 1073 AGTTGTAGAGATAAACCAGCCCTATATAAATCATATATATTTCTCTCCCGCTTTGAATG 1132
QY 763 TCTCGTGTCTCTCTCATCTTTCATCAGCCGCTTTTGAATCTCCGCGAGCTTGACAGAGAAG 822
DB 1133 TCTCGTGTCTCTCTCATCTTTCATCAGCCGCTTTTGAATCTCCGCGAGCTTGACAGAGAAG 1192
QY 823 AACAGGAAGAAGCTAAGAGAGAAGTAAGAGATAATCCAGGAGATTCATTCTCCGTTT 882
DB 1193 AACAGGAAGAAGCTAAGAGAGAAGTAAGAGATAATCCAGGAGATTCATTCTCCGTTT 1252
QY 883 TGAATCTTCTCTCAATCTCATCTTCTTCTTCTCGCTCTTCTTCTTCCAGGTAATAGGAATTT 942
DB 1253 TGAATCTTCTCTCAATCTCA---TCTTCTTCTCGCTCTTCTTCTTCCAGGTAATAGGAATTT 1309
QY 943 TCTGGATCTACTTTATTTTGTGTCGATCTCGATCTTGTGTTTCTCAATTTCTTGTAGACTCG 1002
DB 1310 TCTGGATCTACTTTATTTTGTGTCGATCTCGATCTTGTGTTTCTCAATTTCTTGTAGACTCG 1369
QY 1003 AATTCGTTTAAATTTGGATCTGTGAACCTCCACTTAATCTTTTGGTTTACTAGATCGAT 1062
DB 1370 AATTCGTTTAAATTTGGATCTGTGAACCTCCACTTAATCTTTTGGTTTACTAGATCGAT 1429
QY 1063 CTAAGTTGACCGATCAGTAGCTCGAATATAGCTTACACAGAAATTTGGCTTGACCTTGATGG 1122
DB 1430 CTAAGTTGACCGATCAGTAGCTCGAATATAGCTTACACAGAAATTTGGCTTGACCTTGATGG 1489
QY 1123 AGAGATCCATGTTTCAATGTTACCTGGGAAATGATTTGATATATGTAATGAAATCTGAAT 1182
DB 1490 AGAGATCCATGTTTCAATGTTACCTGGGAAATGATTTGATATATGTAATGAAATCTGAAT 1549
QY 1183 GTTGAAGTTAGATTGAATCTGAACACTGTCAATGTTAGATTGAATCTGAACACTGTTAA 1242
DB 1550 GTTGAAGTTAGATTGAATCTGAACACTGTCAATGTTAGATTGAATCTGAACACTGTTAA 1609
QY 1243 GTTAGATGAAGTTTGTGTATAGATCTTTCGAAACCTTTAGGATTTCTAGTGTCTGATGTTG 1302
DB 1610 GTTAGATGAAGTTTGTGTATAGATCTTTCGAAACCTTTAGGATTTCTAGTGTCTGATGTTG 1669
QY 1303 AACAGAAAGCTATTTCTGATTCATCAACAGGGTTTATTTGACTGTATTTGAACTCTTTTGTG 1362
DB 1670 AACAGAAAGCTATTTCTGATTCATCAACAGGGTTTATTTGACTGTATTTGAACTCTTTTGTG 1729
QY 1363 TGTTTGCAGC 1372
DB 1730 TGTTTGCAGC 1739
RESULT 15
AR438174
LOCUS 1742 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 30 from patent US 6660911.
ACCESSION AR438174
VERSION AR438174.1 GI:40204676
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1742)
AUTHORS Finch, K.L., Plasinski, S. and Wilkinson, J.Q.
TITLE Plant expression constructs
JOURNAL Patent: US 6660911-A 30 09-DEC-2003;
FEATURES
source Location/Qualifiers
1. 1742

/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

```
Query Match      81.3%; Score 1174; DB 6; Length 1742;  
Best Local Similarity 99.3%; Pred. No. 8e-245;  
Matches 1201; Conservative 0; Mismatches 5; Indels 4; Gaps 2;  
  
QY 164 CAACTATTTTATGATGCAAGAGTCAGCATATGTTAAATGATTCAGATTCGTTTGAAC 223  
DB 533 CAACTATTTTATGATGCAAGAGTCAGCATATGTTAAATGATTCAGATTCGTTTGAAC 592  
  
QY 224 GAGTTCGGATGTAGTACTGACCATTTATTAATGTACATATACTAACTCGTGAATGATGATG 283  
DB 593 GAGTTCGGATGTAGTACTGACCATTTATTAATGTACATATACTAACTCGTGAATGATGATG 652  
  
QY 284 ATGAACATTTGATTTCTTTATTTGTTATTAATATCCATAAACAATCATGAAAGACATTTCTT 343  
DB 653 ATGAACATTTGATTTCTTTATTTGTTATTAATATCCATAAACAATCATGAAAGACATTTCTT 712  
  
QY 344 TCAGGGTCTGAATTAATATATGATACAAATCTTAATAGAAAACGAATTAATAGCTTGAAT 403  
DB 713 TCAGGGTCTGAATTAATATATGATACAAATCTTAATAGAAAACGAATTAATAGCTTGAAT 772  
  
QY 404 TGTATGAAATCTAATTCGAACCAACACGACGAGGACTAAAGTTGCCTGGATTTGACT 463  
DB 773 TGTATGAAATCTAATTCGAACCAACACGACGAGGACTAAAGTTGCCTGGATTTGACT 832  
  
QY 464 CGGTTTAAAGTTAAACCACTAAAAAACCAGAGCTGTCTATGTAACACCGGATCGAGCAGGTC 523  
DB 833 CGGTTTAAAGTTAAACCACTAAAAAACCAGAGCTGTCTATGTAACACCGGATCGAGCAGGTC 892  
  
QY 524 ACAGTCATGAGCCTCAAGCAAGCAAGCACTAATCAAGGGGTGAGATGATTAATAGTT 583  
DB 893 ACAGTCATGAGCCTCAAGCAAGCAAGCACTAATCAAGGGGTGAGATGATTAATAGTT 952  
  
QY 584 TAAAAATTTAGTTAAACACGAGGAAAAA-GCTGTCTGACACGAGCTCACGTTATCTTTTACC 642  
DB 953 TAAAAATTTAGTTAAACACGAGGAAAAAAGGCTGTCTGACAGCCAGTCACTTTTACC 1012  
  
QY 643 TGTGTGCAAAATGATTCGTGTCTGCGATTTTAAATTTATTTTGTAAAGGCCGAAAAATAA 702  
DB 1013 TGTGTGCAAAATGATTCGTGTCTGCGATTTTAAATTTATTTTGTAAAGGCCGAAAAATAA 1072  
  
QY 703 AGTGTGAGAGATAAACCAGCTATATAAATTCATATATTTTCTCCCGCTTTGAATTG 762  
DB 1073 AGTGTGAGAGATAAACCAGCTATATAAATTCATATATTTTCTCTCCCGCTTTGAATTG 1132  
  
QY 763 TCTCGTTCTCCTCCTCACCTTCATCAGCCGTTTGAATCTCCGGGACTTGACAGAGAAG 822  
DB 1133 TCTCGTTCTCCTCCTCACCTTCATCAGCCGTTTGAATCTCCGGGACTTGACAGAGAAG 1192  
  
QY 823 AACAGGAGAGAGCTAACAGAGAAAGTAAAGATTAATCCAGGAGATTCATTTCCGTTT 882  
DB 1193 AACAGGAGAGAGCTAACAGAGAAAGTAAAGATTAATCCAGGAGATTCATTTCCGTTT 1252  
  
QY 883 TGAATCTTCTCAATCTCATCTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 942  
DB 1253 TGAATCTTCTCAATCTCA- - -TCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1309  
  
QY 943 TCTGGATCTACTTTATTTGCTGGATCTGATCTTGTGTTTCTCAATTTCTTCTTCTTCTTCTTCT 1002  
DB 1310 TCTGGATCTACTTTATTTGCTGGATCTGATCTTGTGTTTCTCAATTTCTTCTTCTTCTTCTTCT 1369  
  
QY 1003 AATTGTTTAAATTTGGATCTGTGAACCTCCACTAAATCTTTTGGTTTCTTCTTCTTCTTCTTCT 1062  
DB 1370 AATTGTTTAAATTTGGATCTGTGAACCTCCACTAAATCTTTTGGTTTCTTCTTCTTCTTCTTCT 1429  
  
QY 1063 CTAAGTTGACCGATCAGTTAGCTCGATTTATAGCTACCGAATTTGGGCTTGACCTTGATGG 1122  
DB 1430 CTAAGTTGACCGATCAGTTAGCTCGATTTATAGCTACCGAATTTGGGCTTGACCTTGATGG 1489  
  
QY 1123 AGAGATCCATGTTTCATGTTACCTGGAAATGATTTGTATATGTAATGAAATCTGAAC 1182
```

```
DB 1490 AGAGATCCCATGTTTCATGTTTACCTGGGAATGATTTGTATATGTAATGAAATCTGAAC 1549  
QY 1183 GTTGAAGTTAGATTTGAATCTGAACACTGTCAATCTGTAGATTGAATCTGAACACTGTTTAA 1242  
DB 1550 GTTGAAGTTAGATTTGAATCTGAACACTGTCAATCTGTAGATTGAATCTGAACACTGTTTAA 1609  
QY 1243 GTTAGATGAAGTTTGTGTATAGATTCTTCGAAACCTTAGGATTTGTAGTCGTAAGTTG 1302  
DB 1610 GTTAGATGAAGTTTGTGTATAGATTCTTCGAAACCTTTAGGATTTGTAGTCGTAAGTTG 1669  
QY 1303 AACAGAAAGCTATTTCTGATTTCAATCAGGCTTTTATTTGACTGTATTGAACCTTTTGTG 1362  
DB 1670 AACAGAAAGCTATTTCTGATTTCAATCAGGCTTTTATTTGACTGTATTGAACCTTTTGTG 1729  
QY 1363 TGTTCGAGC 1372  
DB 1730 TGTTCGAGC 1739
```

Search completed: March 15, 2004, 00:34:02
Job time : 5821 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2004, 20:42:58 ; Search time 673 Seconds

(without alignments)
9115.005 Million cell updates/sec

Title: US-09-868-744B-1

Perfect score: 1444

Sequence: 1 attagatctcaatacatt.....gtactgaagttagatcc 1444

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapept 1.0

Searched: 3373863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq 29Jan04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002s:*

7: Geneseqn2003as:*

8: Geneseqn2003bs:*

9: Geneseqn2003cs:*

10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1442	99.9	1444	3	AAA61367 Arabidops
2	1341.8	92.9	6385	6	Abk89341 Plasmid p
3	1341.4	92.9	4526	6	Abk89340 Plasmid p
4	1306.2	90.5	11127	6	Aad36967 Arabidops
5	1298.4	89.9	3450	6	Abk89343 Plasmid p
6	1297.4	89.8	3408	6	Abk89342 Plasmid p
7	1297	89.8	2857	6	Abk89344 Plasmid p
8	1174.8	81.4	1219	4	Aad09791 Arabidops
9	1174	81.3	1742	4	Aad09812 ChimERIC
10	1114.6	77.2	12304	7	Abv75876 Luciferas
11	1104.6	76.5	1259	6	Abk52080 Modified
12	1078.8	74.7	15676	3	Aad01288 Arabidops
13	1078.8	74.7	17111	3	Aad01289 Arabidops
14	1078.8	74.7	17116	3	Aad01290 Arabidops
15	1053	72.9	1202	6	Abk52077 Modified
16	1001.8	69.4	1342	6	Abk52081 Modified
17	950.2	65.8	1285	6	Abk52078 Modified
18	784.4	54.3	910	6	Abk52079 Modified
19	732.8	50.7	853	6	Abk52076 Modified
20	579	40.1	1228	6	Abk53111 Transgene
21	475.8	33.0	11461	4	Aad02175 Plasmid p
22	474.8	32.9	12766	4	Aad02174 Plasmid p
23	163	11.3	1486	3	Aac33486 Arabidops

C	24	116.4	8.1	1228	6	ABS53111	Abk53111 Transgene
	25	110	7.6	573	7	ABX56906	Arabidops
	26	94	6.5	94	6	ABK52089	Synthetic
	27	84.4	5.8	97	6	ABK52087	Synthetic
	28	79	5.5	79	6	ABK52084	Synthetic
	29	77.4	5.4	79	6	ABK52092	Synthetic
	30	76	5.3	2000	6	ABZ15497	Arabidops
	31	75.4	5.2	1271	4	AAD09792	Arabidops
	32	73.8	5.1	1800	4	AAD09811	Chimeric
	33	65.4	4.5	67	6	ABK52085	Synthetic
	34	65.4	4.5	74	6	ABK52082	Synthetic
	35	63.4	4.4	77	6	ABK52083	Synthetic
C	36	56.2	3.9	2000	7	ADA71938	Rice gene
	37	51.2	3.5	1134	3	AAc45521	Arabidops
	38	50.8	3.5	1583	3	AAc45216	Arabidops
	39	50.8	3.5	1585	3	AAC33845	Arabidops
	40	50.8	3.5	1612	6	ABN98235	Arabidops
	41	50.4	3.5	6070	6	ABL32240	Human imm
	42	50.4	3.5	6070	6	ABL92198	Chemicall
	43	50.4	3.5	6070	6	ABL49309	Human pol
	44	50.2	3.5	71	6	ABK52088	Synthetic
C	45	49.6	3.4	1940	7	ABX77226	DNA sequ

ALIGNMENTS

RESULT 1

AAA61367

ID AAA61367 standard; DNA; 1444 BP.

XX AC AAA61367;

XX AC AAA61367;

DT 06-NOV-2000 (first entry)

XX Arabidopsis thaliana actin promoter region.

DE Arabidopsis thaliana actin promoter region.

XX Actin 2; ACT2; promoter; oxalate oxidase; OXOX; insecticidal; fungicidal;

KW antiviral; stability; ds.

XX Arabidopsis thaliana.

OS Arabidopsis thaliana.

XX Key Location/Qualifiers

FT primer_bind complement(1..21)

FT primer_bind /tag= a

FT primer_bind 1360..1379

FT primer_bind /tag= b

FT misc_signal 1382..1384

FT misc_signal /tag= c

FT /note= "Translation start site for actin 2"

XX W0200037661-A1.

XX W0200037661-A1.

XX PD 29-JUN-2000.

XX PD 29-JUN-2000.

XX PF 16-DEC-1999; 99WO-GB004317.

XX PF 16-DEC-1999; 99WO-GB004317.

XX PR 21-DEC-1998; 98GB-00028201.

XX PR 21-DEC-1998; 98GB-00028201.

XX (ADVA-) ADVANTA TECHNOLOGY LTD.

XX (ADVA-) ADVANTA TECHNOLOGY LTD.

XX Van Dun CME, Schepers FMA, Pertijs JH;

XX Van Dun CME, Schepers FMA, Pertijs JH;

XX WPI; 2000-442681/38.

XX WPI; 2000-442681/38.

XX Producing recombinant Compositae with increased transformation stability

XX Producing recombinant Compositae with increased transformation stability

XX Claim 2; Fig 3; 22pp; English.

XX Claim 2; Fig 3; 22pp; English.

XX The present sequence is the promoter region and part of the coding

XX sequence of the Arabidopsis gene, actin 2 (ACT2). The promoter may be

XX used in a heterologous DNA construct to drive expression of RNA. This is

XX used in a heterologous DNA construct to drive expression of RNA. This is

XX used in a heterologous DNA construct to drive expression of RNA. This is

XX used in a heterologous DNA construct to drive expression of RNA. This is

PT heavy metal ions, e.g. cadmium, arsenate, useful for phytoremediation of
 PT contaminated soil or water, encodes arsenate reductase coding sequence.
 PS Disclosure; Fig 13B; 131pp; English.

XX The invention relates to a nucleic acid molecule comprising a portion
 CC encoding an arsenate reductase coding sequence and a plant-expressible
 CC transcription regulatory sequence, the coding sequence being operably
 CC linked to the transcription regulatory sequence. The nucleic acid is
 CC useful for producing a plant which is resistant to at least one metal ion
 CC such as a divalent cadmium ion or antimonate, by introducing the nucleic
 CC acid molecule into a plant cell or into plant tissue, selecting for the
 CC presence of the nucleic acid molecule to produce a transgenic plant cell
 CC or plant tissue and regenerating a plant from the transgenic plant cell
 CC or plant tissue. The method further comprises introducing at least one
 CC plant-expressible phytochelatin biosynthetic enzyme coding sequence into
 CC the cell to produce an arsenate metal ion resistant plant. The arsenate
 CC reductase coding sequence is expressed under the control of a plant
 CC promoter which directs expression in the above ground plant part. The
 CC plant is also resistant to cadmium, cobalt, copper, mercury, zinc,
 CC antimony, arsenate and arsenite ions. The transgenic plant is useful for
 CC bioremediation of arsenate and/or cadmium-contaminated environments,
 CC including soil, sediments, mine tailings, water, industrial waste,
 CC soils contaminated with metal ions and for removing and sequestering of
 CC these ions from water, wastewater and aqueous environments. The plants
 CC are also useful for phytoremediation of contaminated soil, sediment,
 CC water and mine tailings. This sequence represents plasmid pACT2B-GUS DNA,
 CC used in the scope of the invention

XX SQ Sequence 6385 BP; 1684 A; 1447 C; 1518 G; 1733 T; 0 U; 3 Other;

Query Match 92.9%; Score 1341.8; DB 6; Length 6385;

Best Local Similarity 98.8%; Pred. No. 4.2e-314;

Matches 1375; Conservative 0; Mismatches 12; Indels 4; Gaps 2;

Qy 1 ATTATGATCTCAAAATACATGATACATATCTCATCTAGATCTAGGTATATATGTAAG 60
 Db 257 ATTATGATCTCAAAATACATGATACATATCTCATCTAGATCTAGGTATATATGTAAG 316
 Qy 61 AAAAGTTTGAGGAATATGNNACGACAAAATGGCTACATCTGATGTAATGGTATCTCAAC 120
 Db 317 AAAAGTTTGAGGAATATGNNACGACAAAATGGCTACATCTGATGTAATGGTATCTCAAC 376
 Qy 121 TCACATATATCTATACCAACATAGTTAGCAAAATTTTACCAACATTTTATGTAAT 180
 Db 377 TCACATATATCTATACCAACATAGTTAGCAAAATTTTACCAACATTTTATGTAAT 436
 Qy 181 GCAAGAGTCAGCATATGTAATTTGATTCAGAAATCGTTTTCAGAGTTCGGATGTAGTAG 240
 Db 437 GCAAGAGTCAGCATATGTAATTTGATTCAGAAATCGTTTTCAGAGTTCGGATGTAGTAG 496
 Qy 241 TAGCCATATTTAATGTACATATCTGTAATGATGATATGATGAAACATTTATCTT 300
 Db 497 TAGCCATATTTAATGTACATATCTGTAATGATGATATGATGAAACATTTATCTT 556
 Qy 301 ATTCTGTAATATCCATAACACATCATGAAAGACATCTTCTTTCAGGGCTGTAATTAAT 360
 Db 557 ATTCTGTAATATCCATAACACATCATGAAAGACATCTTCTTTCAGGGCTGTAATTAAT 616
 Qy 361 TATGATACAAATCTAATAGAAAACGAATTAATTAAGTTGTAATGTAATTAATTAATG 420
 Db 617 TATGATACAAATCTAATAGAAAACGAATTAATTAAGTTGTAATGTAATTAATTAATG 676
 Qy 421 AACAGGCAACACGACGAGACTAAGCTTGGCTGGATGATGATGAAACATTTAACCAC 480
 Db 677 AACAGGCAACACGACGAGACTAAGCTTGGCTGGATGATGATGAAACATTTAACCAC 736
 Qy 481 TAAAAAACCAGGCTGCTATGTAACACGGGATCCAGAGGTACAGTCAATGAGCCATC 540
 Db 737 TAAAAAACCAGGCTGCTATGTAACACGGGATCCAGAGGTACAGTCAATGAGCCATC 796
 Qy 541 AAAGCAAAAGACTAATCAAGGGGTGAGATGATTAATTAATTAATTAATTAATTAATCAAC 600

Db 797 AAGCAAAAGACTAATCAAGGGCTGAGATGATTAATAGTTTAAATTAATTAACAC 856
 Qy 601 GAGGAGAAA-GCTGTCTGACAGCCAGGTCACTTACTTTTACCTGTGGTGAATGATTC 659
 Db 857 GAGGAGAAAAGGCTGTCTGACAGCCAGGTCACTTACTTTTACCTGTGGTGAATGATTC 916
 Qy 660 GTGTCTGTCCATTTAATTAATTTTGAAGGCCGAAATAAGTTGTAAGAGATAAAC 719
 Db 917 GTGTCTGTCCATTTAATTAATTTTGAAGGCCGAAATAAGTTGTAAGAGATAAAC 976
 Qy 720 CGGCTATATAAATTCATATATTTTCTCCCGCTTTTGAATTTGCTCGTTGCTCTCTCA 779
 Db 977 CGGCTATATAAATTCATATATTTTCTCCCGCTTTTGAATTTGCTCGTTGCTCTCTCA 1036
 Qy 780 CTTTATCATCGCCTTTTGAATTTTCTCCCGCTTTGACAGAGAAACAAGAGAGACTA 839
 Db 1037 CTTTATCATCGCCTTTTGAATTTTCTCCCGCTTTGACAGAGAAACAAGAGAGACTA 1096
 Qy 840 AGAGAGAACTAGAGATAATCCAGAGATTCATTTCTCCGTTTGAATTTGCTCTCTCAATCT 899
 Db 1097 AGAGAGAACTAGAGATAATCCAGAGATTCATTTCTCCGTTTGAATTTGCTCTCTCAATCT 1156
 Qy 900 CATCTTTCTTCTCCGCTTTTCTTTTCCAGGTAATAGGAATTTCTGGATCTACTTTAT 959
 Db 1157 CA---TCTTCTTCTCCGCTTTTCTTTTCCAGGTAATAGGAATTTCTGGATCTACTTTAT 1213
 Qy 960 TCTGTGATCTGATCTGTTTCTCAATTTCTTCCATTTCTGATCTGGAATTCGTTAATTGGA 1019
 Db 1214 TCTGTGATCTGATCTGTTTCTCAATTTCTTCCATTTCTGATCTGGAATTCGTTAATTGGA 1273
 Qy 1020 TCTGTGAACCTCCACATAATCTTTTGGTTTACTAGAAATCGATTAAGTTGACCGATCAG 1079
 Db 1274 TCTGTGAACCTCCACATAATCTTTTGGTTTACTAGAAATCGATTAAGTTGACCGATCAG 1333
 Qy 1080 TTAGCTCGATTTAGCTACCAAGATTTGGCTTGACCTTTGATGGAGAGATCCATGTTCAAG 1339
 Db 1334 TTAGCTCGATTTAGCTACCAAGATTTGGCTTGACCTTTGATGGAGAGATCCATGTTCAAG 1393
 Qy 1140 TTACTCTGGAAAATGATTTGATATGTAATTTGAATTTGAATTTGAATTTGAATTTGA 1399
 Db 1394 TTACTCTGGAAAATGATTTGATATGTAATTTGAATTTGAATTTGAATTTGAATTTGA 1453
 Qy 1200 TCTGAACACTGTCAATTTAGATTTAGATTTAGATTTAGATTTAGATTTAGATTTAG 1259
 Db 1454 TCTGAACACTGTCAATTTAGATTTAGATTTAGATTTAGATTTAGATTTAGATTTAG 1513
 Qy 1260 TATAGATTTCTTCAAACTTTAGATTTAGATTTAGATTTAGATTTAGATTTAGATTTCT 1319
 Db 1514 TATAGATTTCTTCAAACTTTAGATTTAGATTTAGATTTAGATTTAGATTTAGATTTCT 1573
 Qy 1320 GATTCATCACTAGGTTTATTTGACCTGATTTGAATCTTTTGGTGTGTTTGCAGCTCATAAA 1379
 Db 1574 GATTCATCACTAGGTTTATTTGACCTGATTTGAATCTTTTGGTGTGTTTGCAGCTCATAAA 1633
 Qy 1380 AAATGGCTGAG 1390
 Db 1634 GGATCCCCGGG 1644

RESULT 3

ABK89340

ID ABK89340 standard; DNA; 4526 BP.

XX AC ABK89340;

XX DT 21-OCT-2002 (first entry)

XX DE Plasmid pACT2B DNA.

XX KW Arsenate reductase; ArsC; cyclic; circular; ds; antimonate; zinc;
 XX KW cadmium; phytochelatin biosynthetic enzyme; arsenite; cobalt;
 XX KW copper; mercury; antimony; soil; sediment; mine tailing; water; air;

KW industrial waste; phytoremediation.
OS Synthetic.
XX WO200248335-A2.
PN 20-JUN-2002.
PD
XX
XX 13-DEC-2001; 2001WO-US048105.
PF
XX
XX 13-DEC-2000; 2000US-0255001P.
PR
XX 22-JUN-2001; 2001US-0300525P.
XX
XX (UYGE-) UNIV GEORGIA RES FOUND INC.
PA
XX
XX Meagher RB, Li Y;
PI
XX
XX WPI; 2002-583507/62.
DR
XX
XX Recombinant DNA molecules for producing transgenic plants tolerant to heavy metal ions, e.g. cadmium, arsenate, useful for phytoremediation of contaminated soil or water, encodes arsenate reductase coding sequence.
PT
PT
PT
XX
PS Disclosure; Fig 12B; 131pp; English.
XX
CC The invention relates to a nucleic acid molecule comprising a portion encoding an arsenate reductase coding sequence and a plant-expressible transcription regulatory sequence, the coding sequence being operably linked to the transcription regulatory sequence. The nucleic acid is useful for producing a plant which is resistant to at least one metal ion such as a divalent cadmium ion or antimonate, by introducing the nucleic acid molecule into a plant cell or into plant tissue, selecting for the presence of the nucleic acid molecule to produce a transgenic plant cell or plant tissue and regenerating a plant from the transgenic plant cell or plant tissue. The method further comprises introducing at least one plant expressible phytochelatin biosynthetic enzyme coding sequence into the cell to produce an arsenite metal ion resistant plant. The arsenate reductase coding sequence is expressed under the control of a plant promoter which directs expression in the above ground plant part. The plant is also resistant to cadmium, cobalt, copper, mercury, zinc, antimony, arsenate and arsenite ions. The transgenic plant is useful for bioremediation of arsenate and/or cadmium-contaminated environments, including soil, sediments, mine tailings, water, industrial waste, ground water and air. The transgenics are also useful for revegetation of soils contaminated with metal ions and for removing and sequestering these ions from water, wastewater and aqueous environments. The plants are also useful for phytoremediation of contaminated soil, sediment, water and mine tailings. This sequence represents plasmid pACT2B DNA, used in the scope of the invention
XX
SQ Sequence 4526 BP; 1203 A; 997 C; 1002 G; 1321 T; 0 U; 3 Other;
Query Match 92.98; Score 1341.4; DB 6; Length 4526;
Best Local Similarity 98.98; Pred. No. 4.2e-314;
Matches 1374; Conservative 0; Mismatches 11; Indels 4; Gaps 2;
QY 1 ATTATGATCTCAATACATATGATATCTCATCTAGATCTAGGTTATCATTTGTAAG 60
DB 257 ATTATGATCTCAATACATATGATATCTCATCTAGATCTAGGTTATCATTTGTAAG 316
QY 61 AAGTTTTGAGGAATATGNNACGAAATGGCTACACTCGATGTAATTTGGTATCTCAAC 120
DB 317 AAGTTTTGAGGAATATGNNACGAAATGGCTAGACTCGATGTAATTTGGTATCTCAAC 376
QY 121 TCAACATATATCTTATACCAACATTTAGTACCAATTTTAAACACTATTTTATGTAT 180
DB 377 TCAACATATATCTTATACCAACATTTAGTACCAATTTTAAACACTATTTTATGTAT 436
QY 181 GCAAGAGTCAGCATATGATTAATTTGATTTAGCAATCGTTTTCAGAGTTCCGATGTAGTAG 240
DB 437 GCAAGAGTCAGCATATGATTAATTTGATTTAGCAATCGTTTTCAGAGTTCCGATGTAGTAG 496
QY 241 TAGCCATTTATTTATGATACATCTAATCGTGAATAGTGAATGATGAAACATTTGATCTT 300

DB 497 TAGCCATTTATTTATGATACATCTAATCGTGAATAGTGAATGAAACATTTGATCTT 556
QY 301 ATTGTATAAATATCCATAAACAACATCATGAAAGACACTTTCTTCAGGGTCTGAATTAAT 360
DB 557 ATTGTATAAATATCCATAAACAACATCATGAAAGACACTTTCTTCAGGGTCTGAATTAAT 616
QY 361 TATGATACAAATTTCTAATAGAAACGAAATTAATTTACGTTGAATTTGATGAATCTTAATG 420
DB 617 TATGATACAAATTTCTAATAGAAACGAAATTAATTTACGTTGAATTTGATGAATCTTAATG 676
QY 421 AACAAAGCCAAACACGACGAGGACTTAACGTTGCTGGATTCGACTCGGTTTAAAGTTAAACCA 480
DB 677 AACAGCCAAACACGACGAGGACTTAACGTTGCTGGATTCGACTCGGTTTAAAGTTAAACCA 736
QY 481 TAAAAAAGCGAGCTGTCATGTAACGCGGATCGAGCAGGTCACAGTCATGAAGCCATC 540
DB 737 TAAAAAAGCGAGCTGTCATGTAACGCGGATCGAGCAGGTCACAGTCATGAAGCCATC 796
QY 541 AAAGCAAAAGAACTAATCCAAAGGGTGAGATGATTAATTTAGTTTAAATTTAGTTAAACAC 600
DB 797 AAAGCAAAAGAACTAATCCAAAGGGTGAGATGATTAATTTAGTTTAAATTTAGTTAAACAC 856
QY 601 GAGGAAAAA-GCTGTCTGACAGCCAGGTCACGTTATCTTTTACCTGTGCTCGAAATGATTC 659
DB 857 GAGGAAAAAGGCTGTCTGACAGCCAGGTCACGTTATCTTTTACCTGTGCTCGAAATGATTC 916
QY 660 GTGTCTGTGATTTAATTAATTTTGAAGGCGGAAATTAAGTTGTAAGAGATAAAC 719
DB 917 GTGTCTGTGATTTAATTAATTTTGAAGGCGGAAATTAAGTTGTAAGAGATAAAC 976
QY 720 CCCCTATATAAATTCATATATTTTCCCTCCCGCTTTGAATTTGCTGCTGCTCTCTCA 779
DB 977 CCCCTATATAAATTCATATATTTTCCCTCCCGCTTTGAATTTGCTGCTGCTCTCTCA 1036
QY 780 CTTTCATCAGCGGTTTGAATTTCCCGGCACTTGACAGAGAAACAAAGGAAGAACTA 839
DB 1037 CTTTCATCAGCGGTTTGAATTTCCCGGCACTTGACAGAGAAACAAAGGAAGAACTA 1096
QY 840 AGAGAAAGTAAGAGATAATCCAGAGATTCATCTCCGTTTGAATTTGAACTCTCTCAATCT 899
DB 1097 AGAGAAAGTAAGAGATAATCCAGAGATTCATCTCCGTTTGAATTTGAACTCTCTCAATCT 1156
QY 900 CATCTCTCTCTCCGCTCTTTCTTCTCCAGGTAATAGAACTTTCTGATCTACTTTAT 959
DB 1157 CA---TCTTCTTCGCTCTTTCTTCCAGGTAATAGAACTTTCTGATCTACTTTAT 1213
QY 960 TGCTGATCTGATCTTGTGTTTCTCAATTTCTTGTGAGATCTCGAAATTCGTTTAAATTTGA 1019
DB 1214 TGCTGATCTGATCTTGTGTTTCTCAATTTCTTGTGAGATCTCGAAATTCGTTTAAATTTGA 1273
QY 1020 TCTGTGAACCTCCACTAAATCTTTTGGTTTCTTACTAGAACTGATCTAGTTGACCGATCAG 1079
DB 1274 TCTGTGAACCTCCACTAAATCTTTTGGTTTCTTACTAGAACTGATCTAGTTGACCGATCAG 1333
QY 1080 TTAGCTCGAATATAGCTTACAGAAATTTGGCTTGAACCTTGATGGAGAGATCCATGTTTCA 1139
DB 1334 TTAGCTCGAATATAGCTTACAGAAATTTGGCTTGAACCTTGATGGAGAGATCCATGTTTCA 1393
QY 1140 TTACCTGGGAAATGATTTGTAATGTAATTTGAATTTGAAATCTGAACTGTTGAGTTAGATGAA 1199
DB 1394 TTACCTGGGAAATGATTTGTAATGTAATTTGAATTTGAAATCTGAACTGTTGAGTTAGATGAA 1453
QY 1200 TCTGACACTGCTCAATCTAGATTTGAATCTGAACTGTTTAAAGTTAGATGAAATTTTGG 1259
DB 1454 TCTGAAACCTGCAATCTTAGATTTGAATCTGAACTGTTTAAAGTTAGATGAAATTTTGG 1513
QY 1260 TATAGATTTCTCGAAACCTTTAGGATTTTGTAGTCTGCTAGCTTGAACAGAAAGCTATTTCT 1319
DB 1514 TATAGATTTCTCGAAACCTTTAGGATTTTGTAGTCTGCTAGCTTGAACAGAAAGCTATTTCT 1573
QY 1320 GATTCAATCAGGGTTTATTTGATCTGATTTGAACTCTTTTGTGTTGTTGAGCTCATAAA 1379

Db 1574 GATTCAATCAGGGTTTATTGACTGTAATTGAACCTTTTGTGTTGTTGAGCTCATAAA 1633
Qy 1380 AAATGGCTG 1388
Db 1634 GGATCCCGG 1642

RESULT 4
AAD36967
ID AAD36967 standard; DNA; 11127 BP.
XX
AC AAD36967;
XX
DT 29-AUG-2003 (revised)
DT 21-AUG-2002 (first entry)
XX
XX Arabidopsis actin-2 promoter-H04 toxin chimeric construct, pZU578.
XX
KW Hybrid delta toxin H04; CryIAb toxin; CryIc toxin; transgenic plant;
KW transgenic; insect control; insect resistance; agricultural; insecticide;
KW actin-2 promoter; pZU578 construct; ds.
XX
OS Bacillus thuringiensis.
OS Arabidopsis sp.
OS Chimeric.

Key Location/Qualifiers
FH promoter 56..1475
FT /*tag= a
FT /note= "Actin-2 promoter"
FT misc_feature 1485..3491
FT /*tag= b
FT /note= "Encodes the toxin portion of H04 and the first 40
FT amino acids of CryIAb tail"
FT promoter 3859..5030
FT /*tag= c
FT /note= "Maize ubiquitin promoter"
FT misc_feature 5052..6271
FT /*tag= d
FT /note= "Phosphomannose isomerase (PMI) marker gene"
XX WO200215701-A2.
XX
XX 28-FEB-2002.
XX
XX 23-AUG-2001; 2001WO-EP009751.
XX
XX 25-AUG-2000; 2000US-0227956P.
XX
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
XX Carozzi NB, Rabe SM, Miles PJ, Warren GW, De Haan PT;
XX WPI; 2002-280838/32.
XX
XX Bacillus thuringiensis hybrid toxin H04 for controlling insects, e.g.
XX fall army worm or European cornborer, and for creating insect resistant
XX plants, comprises domains I and II of CryIAb and domain III of CryIc.
XX
XX Claim 25; Page 127-130; 130pp; English.
XX
XX The present invention relates to Bacillus thuringiensis hybrid delta
XX toxin H04 comprising a N-terminal toxin portion which contain domains I
XX and II from CryIAb toxin joined in the amino to carboxy direction to
XX domain III from CryIc toxin and a C-terminal tail region from a CryIAb
XX toxin. The B. thuringiensis toxins are useful in multiple insect control
XX strategies by inhibiting the ability of insect pests to survive, grow or
XX reproduce or of limiting insect-related damage or loss in crop plants.
XX Nucleotide sequences encoding the toxin are used to generate transgenic
XX plants having resistance to insects. The toxins may be applied in insect
XX infested areas or to prophylactically treat insect susceptible areas or
XX plants to confer protection or resistance against harmful insects. The
XX present DNA sequence is pZU578 chimeric construct comprising Arabidopsis

CC actin-2 promoter operatively linked to the B. thuringiensis H04 toxin
CC portion and the first 40 amino acids of CryIAb tail. (Updated on 29-AUG-
XX 2003 to standardise OS field)
SQ Sequence 11127 BP; 2866 A; 2919 C; 2758 G; 2584 T; 0 U; 0 Other;
Query Match 90.5%; Score 1306.2; DB 6; Length 11127;
Best Local Similarity 98.8%; Pred. No. 28-305;
Matches 1368; Conservative 0; Mismatches 10; Indels 7; Gaps 5;
Qy 1 ATTATGATCTCAAAATACATTTGATACATCTCATCTAGATCTAGGTTATCATTTGTAAG 60
Db |||||
Qy 93 ATTATGATCTCAAAATACATTTGATACATCTCATCTAGATCTAGGTTATCATTTGTAAG 152
Db |||||
Qy 61 AAAGTTTTCAGAAATATGNNACGACAAATGCTGACCTCGATGTAATTGGTATCTCAAC 120
Db |||||
Qy 153 AAAGTTTTCAGAAATATGNNACGACAAATGCTGACCTCGATGTAATTGGTATCTCAAC 212
Db |||||
Qy 121 TCAACATTATCTTATACCAAAACATTAGTTAG-CAAAATTTTAAACAACTA-TTTTATGT 178
Db |||||
Qy 213 TCACATTATCTTATACCAAAACATTAGTTAGCAAAAATTTTAAACAACTATTTTATGT 272
Db |||||
Qy 179 ATGCAAGAGTCAGCATATGTATAATTGATTTCAGAAATCGTTTCACGAGTTCGATGTAGT 238
Db |||||
Qy 273 ATGCAAGAGTCAGCATATGTATAATTGATTTCAGAAATCGTTTCACGAGTTCGATGTAGT 332
Db |||||
Qy 239 AGTAGCCATTATTTAATGTACATACATTAATCGTGAATAGT-ATATGATGAACATTGTAT 297
Db |||||
Qy 333 AGTAGCCATTATTTAATGTACATACATTAATCGTGAATAGTGAATATGATGAACATTGTAT 392
Db |||||
Qy 298 CTTATTGTTATAATATCCATAACACATCATGAAAGACACTTTCTTTTCAGGGTCTGAATT 357
Db |||||
Qy 393 CTTATTGTTATAATATCCATAACACATCATGAAAGACACTTTCTTTTCAGGGTCTGAATT 452
Db |||||
Qy 358 AATTATGATACATTTCTAATAGAAAACGAATTAATTAACGTTGAATTTGATGAATCTAA 417
Db |||||
Qy 453 AATTATGATACATTTCTAATAGAAAACGAATTTAAATTAACGTTGAATTTGATGAATCTAA 512
Db |||||
Qy 418 TTGACACAGCCACACACGACGAGCTAAGCTTGCCTGGATTCATCTCGGTTTAAGTTAAC 477
Db |||||
Qy 513 TTGACACAGCCACACACGACGAGCTAAGCTTGCCTGGATTCATCTCGGTTTAAGTTAAC 572
Db |||||
Qy 478 CACTAAAAAAGCGAGCTGCTCATGTAAACACGCGGATCGAGAGTCCACAGTCATGAAGCC 537
Db |||||
Qy 573 CACTAAAAAAGCGAGCTGCTCATGTAAACACGCGGATCGAGAGTCCACAGTCATGAAGCC 632
Db |||||
Qy 538 ATCAAGCAAAAAGAACTAATCCAGGGGTGAGATGATTAATTAGTTTAAAAATTAGTTAA 597
Db |||||
Qy 633 ATCAAGCAAAAAGAACTAATCCAGGGGTGAGATGATTAATTAGTTTAAAAATTAGTTAA 692
Db |||||
Qy 598 CACGAGGGAAGA-GCTGCTGACAGACGAGTCCAGCTTATCTTTACCTGCTGCGAATGA 656
Db |||||
Qy 693 CACGAGGGAAGAAGGCTGCTGACAGCAGGTCAGCTTATCTTTACCTGCTGCGAATGA 752
Db |||||
Qy 657 TTCGTGCTGCTGATTTTAAATTTTAAATTTTAAAGGCGGAAATTAAGTTTGAAGATA 716
Db |||||
Qy 753 TTCGTGCTGCTGATTTTAAATTTTAAATTTTAAAGGCGGAAATTAAGTTTGAAGATA 812
Db |||||
Qy 717 AACCCGCTATATAATTAATTAATTTTCCCTCCCGCTTTTGAATTTGCTCGTCTCTCC 776
Db |||||
Qy 813 AACCCGCTATATAATTAATTTTCCCTCCCGCTTTTGAATTTGCTCGTCTCTCC 872
Db |||||
Qy 777 TCATTTTCATCAGCGGTTTGAATCTCCGGGCTCTCAGACAGAGAGAAAGAGAGAA 836
Db |||||
Qy 873 TCATTTTCATCAGCGGTTTGAATCTCCGGGCTCTCAGACAGAGAGAAAGAGAGAA 932
Db |||||
Qy 837 CTAAGAGAGAAAGTAAAGAGATAATCCAGAGATTCATTCTCCGTTTGAATTTCTCTCAA 896
Db |||||
Qy 933 CTAAGAGAGAAAGTAAAGAGATAATCCAGAGATTCATTCTCCGTTTGAATTTCTCTCAA 992
Db |||||
Qy 897 TCTCATCTTCTTCTCGCTCTTTCTTCCAGAGTAAAGAGAACTTTCTGGATCTACTTT 956
Db |||||
Qy 993 TCTCA---TCTTCTTCTCGCTCTTTCTTCCAGAGTAAAGAGAACTTTCTGGATCTACTTT 1049
Db |||||

QY 957 ATTGCTGAGCTGATCTGTTTCTCAATTCCTTGAGATCGAATCGTTAAATTT 1016
 Db |||||
 QY 1050 ATTGCTGAGCTGATCTGTTTCTCAATTCCTTGAGATCGAATCGTTAAATTT 1109
 Db |||||
 QY 1017 GGATCTGTGAACCTCCACTAAATCTTTTGGTTTTTACTAGAACTGAATCTTAAGTTGACCGAT 1076
 Db |||||
 QY 1110 GGATCTGTGAACCTCCACTAAATCTTTTGGTTTTTACTAGAACTGAATCTTAAGTTGACCGAT 1169
 Db |||||
 QY 1077 CAGTTAGCTCGATTATAGCTACAGAAATTTGGCTTGACCTTGATGGAGAGATCATGTTTC 1136
 Db |||||
 QY 1170 CAGTTAGCTCGATTATAGCTACAGAAATTTGGCTTGACCTTGATGGAGAGATCATGTTTC 1229
 Db |||||
 QY 1137 ATGTACCTGGGAATGATTTGATATGTGAATTTGAATCTGAATCTGTTGAAGTTAGATT 1196
 Db |||||
 QY 1230 ATGTACCTGGGAATGATTTGATATGTGAATTTGAATCTGAATCTGTTGAAGTTAGATT 1289
 Db |||||
 QY 1197 GAATCTGAACACTGTCATCTGATTTAGATTTGAATCTGAATCTGTTGAAGTTAGATT 1256
 Db |||||
 QY 1290 GAATCTGAACACTGTCATCTGATTTAGATTTGAATCTGAATCTGTTGAAGTTAGATT 1349
 Db |||||
 QY 1257 GTGTATAGATTTCTGAACTTGAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1316
 Db |||||
 QY 1350 GTGTATAGATTTCTGAACTTGAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1409
 Db |||||
 QY 1317 TCTGATTCATCAGGCTTTATTTGACTGTATTTGAACTCTTTTGTGTTGCTGAGCTCAT 1376
 Db |||||
 QY 1410 TCTGATTCATCAGGCTTTATTTGACTGTATTTGAACTCTTTTGTGTTGCTGAGCTCAT 1469
 Db |||||
 QY 1377 AAAAA 1381
 Db 1470 AAAAA 1474

RESULT 5

ID ABK89343 standard; DNA; 3450 BP.
 XX
 AC ABK89343;
 XX
 DT 21-OCT-2002 (first entry)
 XX
 DE Plasmid pECSACT2 DNA.
 XX
 KW Arsenate reductase; ArsC; cyclic; circular; ds; antimonate; zinc;
 KW cadmium; phytochelatin biosynthetic enzyme; arsenite; cobalt;
 KW copper; mercury; antimony; soil; sediment; mine tailing; water; air;
 KW industrial waste; phytoremediation.
 XX
 OS Synthetic.
 XX
 PN WO200248335-A2.
 XX
 PD 20-JUN-2002.
 XX
 PF 13-DEC-2001; 2001WO-US048105.
 XX
 PR 13-DEC-2000; 2000US-0255001P.
 XX
 PR 22-JUN-2001; 2001US-0300525P.
 XX
 PA (UYGE-) UNIV GEORGIA RES FOUND INC.
 XX
 PI Meagher RB, Li Y;
 XX
 DR WPI; 2002-583507/62.
 XX
 PT Recombinant DNA molecules for producing transgenic plants tolerant to
 PT heavy metal ions, e.g. cadmium, arsenate, useful for phytoremediation of
 PT contaminated soil or water, encodes arsenate reductase coding sequence.
 XX
 PS Disclosure; Page 75-77; 131pp; English.
 XX
 CC The invention relates to a nucleic acid molecule comprising a portion

CC encoding an arsenate reductase coding sequence and a plant-expressible
 CC transcription regulatory sequence, the coding sequence being operably
 CC linked to the transcription regulatory sequence. The nucleic acid is
 CC useful for producing a plant which is resistant to at least one metal ion
 CC such as a divalent cadmium ion or antimonate, by introducing the nucleic
 CC acid molecule into a plant cell or into plant tissue, selecting for the
 CC presence of the nucleic acid molecule to produce a transgenic plant cell
 CC or plant tissue and regenerating a plant from the transgenic plant cell
 CC or plant tissue. The method further comprises introducing at least one
 CC plant expressible phytochelatin biosynthetic enzyme coding sequence into
 CC the cell to produce an arsenate metal ion resistant plant. The arsenate
 CC reductase coding sequence is expressed under the control of a plant
 CC promoter which directs expression in the above ground plant part. The
 CC plant is also resistant to cadmium, cobalt, copper, mercury, zinc,
 CC antimony, arsenate and arsenite ions. The transgenic plant is useful for
 CC bioremediation of arsenate and/or cadmium-contaminated environments,
 CC including soil, sediments, mine tailings, water, industrial waste,
 CC groundwater and air. The transgenics are also useful for revegetation of
 CC soils contaminated with metal ions and for removing and sequestering
 CC these ions from water, wastewater and aqueous environments. The plants
 CC are also useful for phytoremediation of contaminated soil, sediment,
 CC water and mine tailings. This sequence represents plasmid pECSACT2 DNA,
 CC used in the scope of the invention
 XX
 SQ Sequence 3450 BP; 946 A; 694 C; 764 G; 1046 T; 0 U; 0 Other;
 Query Match 89.9%; Score 1298.4; DB 6; Length 3450;
 Best Local Similarity 98.1%; Pred. No. 1.1e-303; Indels 8; Gaps 6;
 Matches 1376; Conservative 0; Mismatches 18;
 QY 1 ATTATGATCTCAATACATTTGATACATATCTCATCTAGATCTAGTTCATTTATGTAAG 60
 Db 9 ATGCTGATCTCAATACATTTGATACATATCTCATCTAGATCTAGTTCATTTATGTAAG 68
 QY 61 AAGTTTTGACGAATATGNNACGAAATATGGCTACACTCGATGTAATGGTATCTCAAC 120
 Db 69 AAGTTTTGACGAATATGCGACGACAAATATGGCTAGACTCGATGTAATGGTATCTCAAC 128
 QY 121 TCAACATTTACTATACCAACATTTAGTTAGACAAATTTAAACAACTATTTTTTAGT 178
 Db 129 TCAACATTTACTATACCAACATTTAGTTAGACAAATTTAAACAACTATTTTTTAGT 188
 QY 179 ATGCAAGAGTCAGCATATGTATTAATTTGATTCAGAACTCGTTTTGACGAGTTCGGATGAGT 238
 Db 189 ATGCAAGAGTCAGCATATGTATTAATTTGATTCAGAACTCGTTTTGACGAGTTCGGATGAGT 248
 QY 239 AGTAGCCATTTTAAATGATACATCTAATCTGTAATAGTG-ATATGATGAACAACTGATAT 297
 Db 249 AGTAGCCATTTTAAATGATACATCTAATCTGTAATAGTGATGATGATGAACAACTGATAT 308
 QY 298 CTTATTTGATAAATATCCATAAACAATCATGTAAGACACTTTTCTTCAGGGTCTGAATTT 357
 Db 309 CTTATTTGATAAATATCCATAAACAATCATGTAAGACACTTTTCTTCAGGGTCTGAATTT 368
 QY 358 AATTATGATACAAATTTCTAATAGAAAACGAATTAATTAACGTTGAATTTGATGAATCTAA 417
 Db 369 AATTATGATACAAATTTCTAATAGAAAACGAATTAATTAACGTTGAATTTGATGAATCTAA 428
 QY 418 TTGGAACAGGCCAACCAACGAGGAGCTAATCGTTGCGGATTCGCTGTTAGTTTAC 477
 Db 429 TTGGAACAGGCCAACCAACGAGGAGCTAATCGTTGCGGATTCGCTGTTAGTTTAC 488
 QY 478 CACTAAAAAAGCGAGCTGTCTATGTAACACCGCGGATCGAGCAGGTACAGTCATGAAGCC 537
 Db 489 CACTAAAAAAGCGAGCTGTCTATGTAACACCGCGGATCGAGCAGGTACAGTCATGAAGCC 548
 QY 538 ATCAAGCAAAAAGAACTAAATCCAAAGGGGTGAGATGATTAAATAGTTTAAAAATGATTA 597
 Db 549 ATCAAGCAAAAAGAACTAAATCCAAAGGGGTGAGATGATTAAATAGTTTAAAAATGATTA 608
 QY 598 CACGAGGGAAAA-GCTGCTCTGACAGCCAGGTCAGTTATCTTACCTGTCGTAATGA 656
 Db 609 CACGAGGGAAAAAGGCTGTCTGACAGCCAGGTCAGTTATCTTACCTGTCGTAATGA 668

QY 657 TTCTGTCGTGCGAATTTTAAATTTTGTGAAAGCGCGAAATAAAGTTGTGAAGAGATA 716
 Db 669 TTCTGTCGTGCGAATTTTAAATTTTGTGAAAGCGCGAAATAAAGTTGTGAAGAGATA 728
 QY 717 AACCCGCTATATAAATTCATATATTTTCTCCCGCTTTTGAATTTGTCTGCTGCTGCTCC 776
 Db 729 AACCCGCTATATAAATTCATATATTTTCTCCCGCTTTGAAATTTGTCTGCTGCTGCTCC 788
 QY 777 TCACCTTCATCAGCGGTTTGAATCTCCGCGACTTGCACAGAGAACACAGGAGAGAGA 836
 Db 789 TCACCTTCATCAGCGGTTTGAATCTCCGCGACTTGCACAGAGAACACAGGAGAGAGA 848
 QY 837 CTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATCTCCGCTTTTGAATTTCTCTCAA 896
 Db 849 CTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATCTCCGCTTTTGAATTTCTCTCAA 908
 QY 897 TCTCATCTTCTTCCGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 956
 Db 909 TCTCA--TCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 965
 QY 957 ATTGCTGAGATCTCGATCTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1016
 Db 966 ATTGCTGAGATCTCGATCTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1025
 QY 1017 GGATCTGTGAACCTCCACTAAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1076
 Db 1026 GGATCTGTGAACCTCCACTAAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1085
 QY 1077 CAGTAGCTGATATAGCTACAGAAATTTGCTTGTGAGATCTGGAATTCGTTTAAATTT 1136
 Db 1086 CAGTAGCTGATATAGCTACAGAAATTTGCTTGTGAGATCTGGAATTCGTTTAAATTT 1145
 QY 1137 ATGTTACTGCGGAATGATTTGTATATGTAATCTGAATCTGAATCTGAATCTGAATCTGA 1196
 Db 1146 ATGTTACTGCGGAATGATTTGTATATGTAATCTGAATCTGAATCTGAATCTGAATCTGA 1205
 QY 1197 GAATCTGAACACTGTCAATGTAGATGAATCTGAACACTGTCAATGTAGATGAATCTGA 1255
 Db 1206 GAATCTGAACACTGTCAATGTAGATGAATCTGAACACTGTCAATGTAGATGAATCTGA 1265
 QY 1256 TGCTATAGATTTCTGAAACCTTAGATTTGTAGTGTCTGATGTTGAGAGAGAGATAT 1315
 Db 1266 TGCTATAGATTTCTGAAACCTTAGATTTGTAGTGTCTGATGTTGAGAGAGAGATAT 1325
 QY 1316 TTCTGATTCATCAGGGTTTATTTGACTCTATTGAACTCTTTTGTGTTGTTGCGACTCA 1375
 Db 1326 TTCTGATTCATCAGGGTTTATTTGACTCTATTGAACTCTTTTGTGTTGTTGCGACTCA 1385
 QY 1376 TAAAAAATGGCTGAGGCTGAGC 1397
 Db 1386 TAAACCATGGCAATCCCGAGC 1407

RESULT 6
 ABK9342
 ID ABK9342 standard; DNA; 3408 BP.
 XX
 AC ABK9342;
 XX
 DT 21-OCT-2002 (first entry)
 XX
 DE Plasmid pAtPCSACT2 DNA.
 XX
 KW Arsenate reductase; ArsC; cyclic; circular; ds; antimionate; zinc;
 KW cadmium; phytochelatin biosynthetic enzyme; arsenite; arsenite; cobalt;
 KW copper; mercury; antimony; soil; sediment; mine tailing; water; air;
 XX industrial waste; phytoremediation.
 OS Synthetic.
 OS
 XX W0200248335-A2.
 FN
 XX

PD 20-JUN-2002.
 XX 13-DEC-2001; 2001WO-US048105.
 PF 13-DEC-2001; 2000US-0255001P.
 PR 13-DEC-2001; 2001US-0300525P.
 PR 22-JUN-2001; 2001US-0300525P.
 XX (UYGE-) UNIV GEORGIA RES FOUND INC.
 PA Meagher RB, Li Y;
 PI WPI; 2002-583507/62.
 XX Recombinant DNA molecules for producing transgenic plants tolerant to
 PT heavy metal ions, e.g. cadmium, arsenate, useful for phytoremediation of
 PT contaminated soil or water, encodes arsenate reductase coding sequence.
 XX Disclosure; Page 69-71; 131pp; English.
 CC The invention relates to a nucleic acid molecule comprising a portion
 CC encoding an arsenate reductase coding sequence and a plant-expressible
 CC transcription regulatory sequence, the coding sequence being operably
 CC linked to the transcription regulatory sequence. The nucleic acid is
 CC useful for producing a plant which is resistant to at least one metal ion
 CC such as a divalent cadmium ion or antimonate, by introducing the nucleic
 CC acid molecule into a plant cell or into plant tissue, selecting for the
 CC presence of the nucleic acid molecule to produce a transgenic plant cell
 CC or plant tissue and regenerating a plant from the transgenic plant cell
 CC or plant tissue. The method further comprises introducing at least one
 CC plant-expressible phytochelatin biosynthetic enzyme coding sequence into
 CC the cell to produce an arsenate metal ion resistant plant. The arsenate
 CC reductase coding sequence is expressed under the control of a plant
 CC promoter which directs expression in the above ground plant part. The
 CC plant is also resistant to cadmium, cobalt, copper, mercury, zinc,
 CC antimony, arsenate and arsenite ions. The transgenic plant is useful for
 CC bioremediation of arsenate and/or cadmium-contaminated environments,
 CC including soil, sediments, mine tailings, water, industrial waste,
 CC groundwater and air. The transgenics are also useful for revegetating
 CC soils contaminated with metal ions and for removing and sequestering
 CC these ions from water, wastewater and aqueous environments. The plants
 CC are also useful for phytoremediation of contaminated soil, sediment,
 CC water and mine tailings. This sequence represents plasmid pAtPCSACT2 DNA,
 CC used in the scope of the invention
 XX Sequence 3408 BP; 980 A; 638 C; 693 G; 1097 T; 0 U; 0 Other;

Query Match 89.8%; Score 1297.4; DB 6; Length 3408;
 Best Local Similarity 98.5%; Pred. No. 1.9e-303;
 Matches 1372; Conservative 0; Mismatches 13; Indels 8; Gaps 6;
 QY 1 ATTATGATCTCAAAATACATTCATACATATCTCATCTAGATCTAGGTTATCATTTAAG 60
 Db 9 ATGCTGATCTCAAAATACATTCATACATATCTCATCTAGATCTAGGTTATCATTTAAG 68
 QY 61 AAAGTTTGGAGAAATATGNNACGACAAATGGCTACACTCGATGTAATGGTAATCTCAAC 120
 Db 69 AAAGTTTGGAGAAATATGNNACGACAAATGGCTACACTCGATGTAATGGTAATCTCAAC 128
 QY 121 TCAACATTTACTTATACCAACATTTAGTTAG-CAAAATTTTAAACACTA-TTTTATGT 178
 Db 129 TCAACATTTACTTATACCAACATTTAGTTAGCAAAATTTTAAACACTA-TTTTATGT 188
 QY 179 ATGCAAGAGTCAGCATATGTAATTTGATTCAGATCGTTTTCAGGAGTTCCGATGAT 238
 Db 189 ATGCAAGAGTCAGCATATGTAATTTGATTCAGATCGTTTTCAGGAGTTCCGATGAT 248
 QY 239 AGTAGCAATTTTATGTAATTTGATTCAGATCGTTTTCAGGAGTTCCGATGAT 297
 Db 249 AGTAGCAATTTTATGTAATTTGATTCAGATCGTTTTCAGGAGTTCCGATGAT 308
 QY 298 CTTATTTGTAATAATATCCATAAACAACATCATCAAAAGACATTTCTTTTCCAGGTCGAAT 357
 Db 309 CTTATTTGTAATAATATCCATAAACAACATCATCAAAAGACATTTCTTTTCCAGGTCGAAT 368

QY 61 AAAGTTTTCAGCAATATGNNACGACAAATGGCTACACTCGATGTAATTTGGTTACTCAAC 120
Db 69 AAAGTTTTCAGCAATATGGCAGCAAAATGGCTAGACTCGATGTAATTTGGTTACTCAAC 128
QY 121 TCAACATTATACCTTATACCAACAACTTAGTTAG-CAAAATTTAAACAACCTA-TTTTATGT 178
Db 129 TCAACATTATACCTTATACCAACAACTTAGTTAGCAAAATTTAAACAACCTA-TTTTATGT 188
QY 179 ATGCAAGATCAGCATATGTAATTCGATTCGAAATCGTTTTCAGAGTTCGGATAGT 238
Db 189 ATGCAAGATCAGCATATGTAATTCGATTCGAAATCGTTTTCAGAGTTCGGATAGT 248
QY 239 AGTAGCCATTATTTAATGATACATACTAATTCGTAATGCG-ATATCATGCAACCACTTCTAT 297
Db 249 AGTAGCCATTATTTAATGATACATACTAATTCGTAATGCGTGAATGATGATGAAACATTTGAT 308
QY 298 CTTATTCGTAATATCCATAAACAACATCATGAAAGACATTTCTTTTCAGGCTCGAAT 357
Db 309 CTTATTCGTAATATCCATAAACAACATCATGAAAGACATTTCTTTTCAGGCTCGAAT 368
QY 358 AATTATGATACATTTCTAATAGAAACGAATTAATTTACGTTGAAATGTAATGAAATCTAA 417
Db 369 AATTATGATACATTTCTAATAGAAACGAATTAATTTACGTTGAAATGTAATGAAATCTAA 428
QY 418 TTGAACAAGCCAAACACGACGAGGACTTAACGTTGCTGGATTGACTCGGTTTAAAGTTAAC 477
Db 429 TTGAACAAGCCAAACACGACGAGGACTTAACGTTGCTGGATTGACTCGGTTTAAAGTTAAC 488
QY 478 CACTAAAAAAGCGAGCTGTGATGTAACACGCGGATCGAGGCTCACAGTCATGAAGCC 537
Db 489 CACTAAAAAAGCGAGCTGTGATGTAACACGCGGATCGAGGCTCACAGTCATGAAGCC 548
QY 538 ATCAACGAACCACTAATCCAGGGGTGAGATGATTAATTTAGTTTAAATTTAGTTAA 597
Db 549 ATCAACGAACCACTAATCCAGGGGTGAGATGATTAATTTAGTTTAAATTTAGTTAA 608
QY 598 CACGAGGGAAGA-GCTGTCTGACAGCCAGGTCACTGTTATCTTTACCTGTGGTGGAAATGA 656
Db 609 CACGAGGGAAGAAGCTGTCTGACAGCCAGGTCACTGTTATCTTTACCTGTGGTGGAAATGA 668
QY 657 TTCGTGCTGTGATTTAATTTATTTTGAAGGCGGAAATTAAGTTGTAAGAGATA 716
Db 669 TTCGTGCTGTGATTTAATTTATTTTGAAGGCGGAAATTAAGTTGTAAGAGATA 728
QY 717 AACCGGCTATATAATTCATATATTTCTCCCGCTTTGAAATTCCTGCTGCTCTCC 776
Db 729 AACCGGCTATATAATTCATATATTTCTCCCGCTTTGAAATTCCTGCTGCTCTCC 788
QY 777 TCACCTTTTCATCAGCGGTTTGAATCTCCGCGACTTGACAGAGAACAAGGAAGA 836
Db 789 TCACCTTTTCATCAGCGGTTTGAATCTCCGCGACTTGACAGAGAACAAGGAAGA 848
QY 837 CTAAGAGAGAAGTAAGAGATAATCCAGAGATTCATTTCCGTTTGAATCTTCCTCAA 896
Db 849 CTAAGAGAGAAGTAAGAGATAATCCAGAGATTCATTTCCGTTTGAATCTTCCTCAA 908
QY 897 TCTCATCTCTCTTCTCCGCTCTTTCTTCCAGGTAATAGCACTTTCTGGATCTACTTT 956
Db 909 TCTCA-CTCTCTCTCCGCTCTTTCTTCCAGGTAATAGCACTTTCTGGATCTACTTT 965
QY 957 ATTTGCTGGATCTCGATCTTTCTCAATTTCTTGAGATCTGGAATTCGTTAAATTT 1016
Db 966 ATTTGCTGGATCTCGATCTTTCTCAATTTCTTGAGATCTGGAATTCGTTAAATTT 1025
QY 1017 GGAATCTGGAACCTCCATAATCTTTTGGTTTACTAGAAATCGATCAAGTCAACGAT 1076
Db 1026 GGAATCTGGAACCTCCATAATCTTTTGGTTTACTAGAAATCGATCAAGTCAACGAT 1085
QY 1077 CAGTTAGCTCGATTATAGTACCAAGATTTGGCTTGACCTTGATGAGAGATCCATCTTC 1136
Db 1086 CAGTTAGCTCGATTATAGTACCAAGATTTGGCTTGACCTTGATGAGAGATCCATCTTC 1145

QY 1137 ATGTTACCTCGAAATGATTGTTGATATGCAATGAATCTGAACCTGTTGAAGTTAGATT 1196
Db 1146 AIGTTTACCTCGAAATGATTGTTGATATGCAATGAATCTGAACCTGTTGAAGTTAGATT 1205
QY 1197 GAATCTGAACACTGTCAATGTTAGATTGAATCTGAACACTGTTTAA-GTTAGATGAAGTT 1255
Db 1206 GAATCTGAACACTGTCAATGTTAGATTGAATCTGAACACTGTTTAAAGTTTGAATGAAGTT 1265
QY 1256 TGTGTATAGATTCTTCGAAACCTTAGGATTTGTAGTGTGCTGCTGAACAGAAAGCTAT 1315
Db 1266 TGTGTATAGATTCTTCGAAACCTTAGGATTTGTAGTGTGCTGCTGAACAGAAAGCTAT 1325
QY 1316 TTCTGATTCATCAGGCTTTATTTGACTGTATTGAACCTTTTTTGTGTTTGCAGCTCA 1375
Db 1326 TTCTGATTCATCAGGCTTTATTTGACTGTATTGAACCTTTTTTGTGTTTGCAGCTCA 1385
QY 1376 TAAAAAATGGC 1386
Db 1386 TAAACCATGGC 1396
RESULT 8
AAD09791
ID AAD09791 standard; DNA; 1219 BP.
XX
AC AAD09791;
XX
DT 10-SEP-2001 (first entry)
XX
DE Arabidopsis thaliana Act2 promoter including intron of Act2 gene.
XX
KW DNA construct; structural gene; SG; non-translated region; NTR; actin;
KW Act; elongation factor 1 alpha; EF1alpha; weed control; EPSP synthase;
KW glyphosate tolerance; glyphosate oxidoreductase; herbicide tolerance;
KW PCR primer; Act2 promoter; ds.
XX
OS Arabidopsis thaliana.
XX
FH Key Location/Qualifiers
FT promoter 1..764
FT /*tag= a
FT /note= "Act2 promoter"
FT misc_signal 597
FT /*tag= b
FT /note= "Transcription start site"
FT intron 765..1215
FT /*tag= c
XX
PN WO200144457-A2.
XX
PD 21-JUN-2001.
XX
PF 12-DEC-2000; 2000WO-US033633.
XX
PR 16-DEC-1999; 99US-0171173P.
XX
PA (MONS) MONSANTO CO.
XX
PI Fincher KL, Flasiniski S, Wilkinson JQ;
XX
DR WPI; 2001-408480/43.
XX
PT Novel DNA construct for controlling weeds, containing Arabidopsis actin
PT promoter sequences and elongation factor 1-alpha promoter sequences
PT operably linked to structural genes that function in crop plants.
XX
PS Claim 1; Page 85; 101pp; English.
XX
CC The present invention relates to DNA construct comprising an expression
CC cassette containing a promoter DNA sequence operably linked to a
CC structural gene (SG) encoding an agronomically useful protein, and a 3'
CC non-translated region (NTR) that functions in plants to cause the
CC addition of polyadenylated nucleotides to the 3' end of the RNA sequence,

CC operably linked to SG. The present invention further relates to plant
CC expression constructs that comprise Arabidopsis actin (Act) promoter
CC sequences Actia, Actb, Act2, Act3, Act7, Act11, Act12, elongation
CC factor 1 alpha (E1alpha) promoter sequence, fragments and cis elements
CC derived from these promoters operably linked to heterologous structural
CC gene sequences that function in crop plant cells. The DNA constructs are
CC useful for controlling weeds, by providing a crop plant transformed with
CC DNA construct comprising promoter DNA sequence that is functional in a
CC plant cell, a glyphosate tolerance gene (which is an EPSP synthase gene
CC or a glyphosate oxidoreductase gene) or a herbicide tolerance gene. The
CC present sequence is Arabidopsis thaliana Act2 promoter including intron
CC of Act2 gene
XX

SQ Sequence 1219 BP; 370 A; 207 C; 233 G; 409 T; 0 U; 0 Other;

Query Match 81.4%; Score 1174.8; DB 4; Length 1219;
Best Local Similarity 98.7%; Pred. No. 5.9e-274;
Matches 1206; Conservative 0; Mismatches 12; Indels 4; Gaps 2;

QY 164 CAACATATTTTATGATGCAAGAGTCAGCATATGTATAATTTGATTCAGAACTCGTTTGAC 223
DB 1 CAACATATTTTATGATGCAAGAGTCAGCATATGTATAATTTGATTCAGAACTCGTTTGAC 60
QY 224 GAGTTCGAGTAGTAGGACCAATTAATTTAATGTACATCAATTAATGTGAATAGTAGTATG 283
DB 61 GAGTTCGAGTAGTAGGACCAATTAATTTAATGTACATCAATTAATGTGAATAGTAGTATG 120
QY 284 ATGAACATGATGATCTTATTTGATTAATATCCATTAACACATCATGAAGACATTTCTT 343
DB 121 ATGAACATGATGATCTTATTTGATTAATATCCATTAACACATCATGAAGACATTTCTT 180
QY 344 TCAGGGTCTGAATTAATTAATGATACAAATTTAAATAGAAAACGAATTAATTAATCACTTGAAT 403
DB 181 TCAGGGTCTGAATTAATTAATGATACAAATTTAAATAGAAAACGAATTAATTAATCACTTGAAT 240
QY 404 TGTATGAATCTTAATTAAGAACAGCAACCAACGACGAGCACTAACGTGCTGAGTATGACT 463
DB 241 TGTATGAATCTTAATTAAGAACAGCAACCAACGACGAGCACTAACGTGCTGAGTATGACT 300
QY 464 CGGTTTAAGTTAACCACTAAATAACCGAGAGCTGCTAATCAACGCGGATCGACAGTTC 523
DB 301 CGGTTTAAGTTAACCACTAAATAACCGAGAGCTGCTAATCAACGCGGATCGACAGTTC 360
QY 524 ACAGTCAATGAAGCCATCAAGCAAAAAGAACTAATCAAGGGGTGAGATGATTAATAGTT 583
DB 361 ACAGTCAATGAAGCCATCAAGCAAAAAGAACTAATCAAGGGGTGAGATGATTAATAGTT 420
QY 584 TAAAAATTAGTTAAACAGAGGAAAAA-GCTGTCTGACAGCGAGTCACTTTATCTTTTACC 642
DB 421 TAAAAATTAGTTAAACAGAGGAAAAAAGGCTGTCTGACAGCGAGTCACTTTATCTTTTACC 480
QY 643 TGTGTCGAAATGATTCGTGCTGATTTTAAATTAATTTTAAAGGCGGAAATAA 702
DB 481 TGTGTCGAAATGATTCGTGCTGATTTTAAATTAATTTTAAAGGCGGAAATAA 540
QY 703 AGTTGTAAGAGATAAACCAGGCTATATAATTAATTTTCTCCCGCTTTGAATG 762
DB 541 AGTTGTAAGAGATAAACCAGGCTATATAATTAATTTTCTCCCGCTTTGAATG 600
QY 763 TCTGTTGTCCTCTCACTTTCACTCAGCGGTTTGAATCTCCGCGACTTTCAGAGAGAG 822
DB 601 TCTGTTGTCCTCTCACTTTCACTCAGCGGTTTGAATCTCCGCGACTTTCAGAGAGAG 660
QY 823 AACAGGAAGAAGACTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTTCTCGTTT 882
DB 661 AACAGGAAGAAGACTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATTTCTCGTTT 720
QY 883 TGAATCTTCTCAATCTCATCTTCTTCTTCCGCTCTTTCTTCCAGAGTAATAGAACTT 942
DB 721 TGAATCTTCTCAATCTCA---TCTTCTTCCGCTCTTTCTTCCAGAGTAATAGAACTT 777
QY 943 TCTGATCTACTTATTTTGTGCTGGATCTCGATCTGTTTCTCAATTTCTTGGATCTGG 1002

DB 778 TCTGATCTACTTATTTTGTGCTGGATCTCGATCTGTTTCTCAATTTCTTGGATCTGG 837
QY 1003 AATTCGTTTAAATTTGATCTGTGAACCTCCACTAAATCTTTTGGTTTACTAGAACTGAT 1062
DB 838 AATTCGTTTAAATTTGATCTGTGAACCTCCACTAAATCTTTTGGTTTACTAGAACTGAT 897
QY 1063 CTAAAGTTGACCGATCAGTTAGCTGATATAGCTACCAAGAAATTTGGCTTGACCTTGATGG 1122
DB 898 CTAAAGTTGACCGATCAGTTAGCTGATATAGCTACCAAGAAATTTGGCTTGACCTTGATGG 957
QY 1123 AGAGATCCATGTTTCAATGTTACCTGGGAAATGATTTGTATGTGAATTTGAATCTGAACT 1182
DB 958 AGAGATCCATGTTTCAATGTTACCTGGGAAATGATTTGTATGTGAATTTGAATCTGAACT 1017
QY 1183 GTTGAAGTTAGATGAATCTGAACACTGTCTAGATTTAGATTTGAATCTGAACACTGTTTAA 1242
DB 1018 GTTGAAGTTAGATGAATCTGAACACTGTCTAGATTTAGATTTGAATCTGAACACTGTTTAA 1077
QY 1243 GTTGAAGTTAGATTTGTGTATAGATTTCTTGAACCTTAGATTTGTGTCTGACCTGTTG 1302
DB 1078 GTTGAAGTTAGATTTGTGTATAGATTTCTTGAACCTTAGATTTGTGTCTGACCTGTTG 1137
QY 1303 AACAGAAAGCTATTTCTGATTCATCAATCAGGGTTTATTTGACGTGATTTGAATCTTTTGTG 1362
DB 1138 AACAGAAAGCTATTTCTGATTCATCAATCAGGGTTTATTTGACGTGATTTGAATCTTTTGTG 1197
QY 1363 TGTGTCGAGCTCATAAAATG 1384
DB 1198 TGTGTCGAGGAGCTCACCATG 1219
RESULT 9
AAD09812
ID AAD09812 standard; DNA; 1742 BP.
XX
AC AAD09812;
XX
DT 11-SEP-2003 (revised)
DT 10-SEP-2001 (first entry)
XX
DE Chimeric CamV-Act2 promoter including first intron of Act2 gene.
KW DNA construct; structural gene; SG; non-translated region; NTR; actin;
KW Act; elongation factor 1 alpha; E1alpha; weed control; EPSP synthase;
KW glyphosate tolerance; glyphosate oxidoreductase; herbicide tolerance;
KW PCR primer; Act2 promoter; Cauliflower mosaic virus; CamV; chimeric; ds.
XX
OS Cauliflower mosaic virus.
OS Arabidopsis thaliana.
OS Chimeric.
PH Key Location/Qualifiers
FT promoter 1..523 a
FT /note= "CamV promoter"
FT misc_feature 534..1742
FT /tag= b
FT /note= "Arabidopsis thaliana Act2 promoter, intron and 5'
FT UTR region"
XX
PN W020014457-A2.
XX
PD 21-JUN-2001.
XX
PF 12-DEC-2000; 2000WO-US033633.
XX
PR 16-DEC-1999; 99US-0171173P.
XX (MONS) MONSANTO CO.
XX Fincher KL, Flasiniski S, Wilkinson JQ;
XX MPI; 2001-408480/43.
DR

CDS 1243..2895
 /*tag= d
 /product= "Luciferase"
 polyA_site 2929..3570
 /*tag= e
 /note= "RbcS E9 polyA region"
 WO200281647-A2.
 17-OCT-2002.
 08-APR-2002; 2002WO-US011116.
 06-APR-2001; 2001US-0282094P.
 (SRI) SRIPPS RES INST.
 Kay SA, Kuhlmann T, Lerner RA;
 WPI; 2003-058526/05.
 Novel genetically modified plant cell, useful as research tool, comprises a heterologous nucleotide sequence encoding a bioluminescent polypeptide expressed in an amount sufficient to produce visible light.
 Claim 21; Page 90-96; 96pp; English.
 The present sequence is that of luciferase expression vector ACT-OM-LUC comprising an expression construct composed of an actin 2 regulatory element comprising an enhancer and promoter, a tobacco mosaic virus omega translational enhancer, a nucleotide sequence encoding luciferase, and an RbcS E9 polyA region. Claimed vectors, including ACT-OM-LUC, can be used to transform a plant cell such that the cell expresses luciferase in an amount sufficient to produce at least 750,000 photons of visible light/sq mm/second. Also claimed are visibly bioluminescent transgenic plants that contain the genetically modified plant cell. The transgenic plant may be a monocot or dicot, including an angiosperm, cereal, legume, oilseed plant or hardwood tree, or an ornamental plant such as petunia or carnation (all claimed). The transgenic plants are useful as research tools and have ornamental value
 Sequence 12304 BP; 2928 A; 3115 C; 3142 G; 3119 T; 0 U; 0 Other;
 Query Match 77.2%; Score 1114.6; DB 7; Length 12304;
 Best Local Similarity 98.0%; Pred. No. 4.3e-259;
 Matches 1150; Conservative 0; Mismatches 19; Indels 4; Gaps 2;
 QY 224 GAGTTCGGATGTAGTAGTACCAATTTTAAATGATACATCTAATCGTAATAGTATG 283
 DB 1 GATCCCGGATGTAGTAGTACCAATTTTAAATGATACATCTAATCGTAATAGTATG 60
 QY 284 ATGAACATGTATCTTATTGTATATAATCCATAACACATCATGAAGACACTTTCTT 343
 DB 61 ATGAACATGTATCTTATTGTATATAATCCATAACACATCATGAAGACACTTTCTT 120
 QY 344 TCAGGCTGTGAATTAATATGATGATCAATTTCTAATAGAAAACGAATTAATGCTTGAAT 403
 DB 121 TCAGGCTGTGAATTAATATGATGATCAATTTCTAATAGAAAACGAATTAATGCTTGAAT 180
 QY 404 TGATGAACTTAATTTGAACAGCCACCAACGACGAGACTACGTTGCCGTGATTCAT 463
 DB 181 TGATGAACTTAATTTGAACAGCCACCAACGACGAGACTACGTTGCCGTGATTCAT 240
 QY 464 CGGTTTAAGTTAACCACTTAAAAACGGAGCTGTCTGATTAACACGGGATCGACAGGTC 523
 DB 241 CGGTTTAAGTTAACCACTTAAAAACGGAGCTGTCTGATTAACACGGGATCGACAGGTC 300
 QY 524 ACAGTCATGAGCCATCAAGCAAAAGAACTAATCCAAAGGGGTGAGATGATTAATAGTT 583
 DB 301 ACAGTCATGAGCCATCAAGCAAAAGAACTAATCCAAAGGGGTGAGATGATTAATAGTT 360
 QY 584 TAAATATAGTTAAACACGGGAAA- GCTGTCTGACGACGAGTCACTTATCTTTACC 642

Db 361 TAAAAATTAGTTAAACAACAGAGGAAAGGCTGTCTGACAGCCAGGTCACTTATCTTTACC 420
 QY 643 TGTGTCGAAAATGATTCGTGTCTGTGATTTTAAATTTTGTGAAAGCCGGAATAA 702
 Db 421 TGTGTCGAAAATGATTCGTGTCTGTGATTTTAAATTTTGTGAAAGCCGGAATAA 480
 QY 703 AGTTGTAAAGAGATAAACCCGCTATATAAATTCATATATTTCTCCCGCTTTTGAATTG 762
 Db 481 AGTTGTAAAGAGATAAACCCGCTATATAAATTCATATATTTCTCTCCGCTTTGAATTG 540
 QY 763 TCTCGTTGCTCCCTCACTTTTCATCAGCCGTTTGAATTCCTCCGCGACTTGACAGAGAAG 822
 Db 541 TCTCGTTGCTCCCTCACTTTTCATCAGCCGTTTGAATTCCTCCGCGACTTGACAGAGAAG 600
 QY 823 AACAGGAGAGAGACTAAGAGAGAAAGTAAGAGATAATCCAGAGATTCATCTCCGCTT 882
 Db 601 AACAGGAGAGAGACTAAGAGAGAAAGTAAGAGATAATCCAGAGATTCATCTCCGCTT 660
 QY 883 TGAATCTTCTCAATCTCATCTTCTTCTTCCGCTCTTCTTCTTCCAAAGGTAATAGAACTT 942
 Db 661 TGAATCTTCTCAATCTCACTCA---TCTTCTCCGCTCTTCTTCTTCCAAAGGTAATAGAACTT 717
 QY 943 TCTGATCTACTTTATTTGCTGATCTCGATCTGTCTTCTCAATTTCTTCTGAGATCTGG 1002
 Db 718 TCTGATCTACTTTATTTGCTGATCTCGATCTGTCTTCTCAATTTCTTCTGAGATCTGG 777
 QY 1003 AATTCGTTTAAATTTGGATCTGTGAACCTCCACTAAATCTTTTGGTTTACTAGAACTGAT 1062
 Db 778 AATTCGTTTAAATTTGGATCTGTGAACCTCCACTAAATCTTTTGGTTTACTAGAACTGAT 837
 QY 1063 CTAAGTTGACCGATCAGTTAGCTCGATATATAGCTACAGAAATTTGGCTTGACCTTGATGG 1122
 Db 838 CTAAGTTGACCGATCAGTTAGCTCGATATATAGCTACAGAAATTTGGCTTGACCTTGATGG 897
 QY 1123 AGATCCATCTGTTCTGTTTACCTGGGAATGATTTGTATGTATGTGAATTCGAATCTGAACT 1182
 Db 898 AGATCCATCTGTTCTGTTTACCTGGGAATGATTTGTATGTATGTGAATTCGAATCTGAACT 957
 QY 1183 GTTCAAGTTAGATTAATCTGAACATCTGAACATCTGAATGTAGATTAATCTGAACATCTGTTAA 1242
 Db 958 GTTCAAGTTAGATTAATCTGAACATCTGAACATCTGAATGTAGATTAATCTGAACATCTGTTAA 1017
 QY 1243 GTTCAAGTTAGATTAATCTGAACATCTGAACATCTGAATGTAGATTAATCTGAACATCTGTTAA 1302
 Db 1018 GTTCAAGTTAGATTAATCTGAACATCTGAACATCTGAATGTAGATTAATCTGAACATCTGTTAA 1077
 QY 1303 AACAGAAAGCTATTTCTGATTCATCAATCAGGGTTTATTTGACTGTATTTGAACCTTTTGTG 1362
 Db 1078 AACAGAAAGCTATTTCTGATTCATCAATCAGGGTTTATTTGACTGTATTTGAACCTTTTGTG 1137
 QY 1363 TGTTCGAGCTCATAAAAAATGGCTGAGGCTGA 1395
 Db 1138 TGTTCGAGCTCAGGATCCATCGATTAAGCTTTA 1170
 RESULT 11
 ABKS2080
 ID ABKS2080 standard; DNA; 1259 BP.
 XX ABKS2080;
 AC ABKS2080;
 DT 13-AUG-2002 (first entry)
 XX Modified plant promoter #5.
 DE Promoter; ds: plant; soybean; peanut; alfalfa; tomato; eggplant; potato;
 KW cabbage; turnip; rapeseed; apple; pear; berry; cucumber; carrots;
 KW transgenic plant.
 XX Synthetic.
 OS US2002049992-A1.
 FN US2002049992-A1.
 XX

KW Arabidopsis transformation binary vector; pAct2-bin; gene expression;
KW transgenic organism; ds.
XX Synthetic.
XX WO200032800-A1.
XX 08-JUN-2000.
XX 30-NOV-1999; 99WO-US028123.
XX 01-DEC-1998; 98US-0110437P.
XX (DOWC) DOW AGROSCIENCES LLC.
XX Van Der Geest AHM, Ainley WM, Cowen NM, Welter ME, Woosley AT;
XX WPI; 2000-412345/35.
XX An isolated DNA molecule for use as a matrix attachment region to
PT increase expression of genes introduced in transformed plants comprises a
PT 298 base pair sequence described in the specification.
XX Example 3; Page 55-59; 73pp; English.
XX The patent discloses a DNA molecule, useful as matrix attachment region
CC (MAR) or scaffold attachment region, to increase the expression of genes
CC introduced in transformed plants. MARs are located in non-transcribed
CC regions of genes and form the physical boundaries of individual DNA
CC loops. They are rich in adenosine and thymine bases and contain certain
CC conserved sequence elements and structural features. They are about 300-
CC 2000 bp in length. Increased levels of expression of DNA introduced into
CC plants can be achieved by use of MAR. They can also reduce the position
CC effect in transgenic organisms. The present DNA sequence is the
CC Arabidopsis transformation binary vector pAct2-bin, that contains a
CC Act2/GUS/nos cassette (Act2 transcription initiation region/GUS
CC structural gene/nos 3' untranslated region), 19S/NPTII/orf25/polyA as a
CC selectable marker and 35S/GFP/nos as an independent reporter gene. This
CC binary vector is used to construct other vectors, used to test the two
CC orientations of the artificial MAR dimers in Arabidopsis
XX
SQ Sequence 15676 BP; 4226 A; 3741 C; 3838 G; 3868 T; 0 U; 3 Other;
Query Match 74.7%; Score 1078.8; DB 3; Length 15676;
Best Local Similarity 98.3%; Pred. No. 2.1e-250;
Matches 1154; Conservative 0; Mismatches 12; Indels 8; Gaps 6;
QY 205 GATTCAGATCGTTTGGACGAGTTCGGATGTAGTAGTACCCATTATTTAATGTACATACT 264
DB 84 GAATTCGGCTTGTGTTGACGAGTTCGGATGTAGTAGTACCCATTATTTAATGTACATACT 143
QY 265 AATCGTGAATAGTG-ATATGATGAACATTGTATCTATTGTATATAAATATCCATAAACAC 323
DB 144 AATCGTGAATAGTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 203
QY 324 ATCATGAAGACACATTTCTTTTACGAGTTCGATTAATATATGATGATGATGATGATGATGATGAT 383
DB 204 ATCATGAAGACACATTTCTTTTACGAGTTCGATTAATATATGATGATGATGATGATGATGATGAT 263
QY 384 CGAATTAATATAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 443
DB 264 CGAATTAATATAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 323
QY 444 TAAGGTGTCCTGGATTGACTCGGTTTAAAGTTAACCACTAAAAAAGCGAGTGTGATGTA 503
DB 324 TAAAGTGTGCTGGATTGACTCGGTTTAAAGTTAACCACTAAAAAAGCGAGTGTGATGTA 383
QY 504 ACACGCGGATTCGACAGGTACAGTATGAAGCCATCAAGCAAAAGAAAGAACTAATCCAAAG 563
DB 384 ACACGCGGATTCGACAGGTACAGTATGAAGCCATCAAGCAAAAGAAAGAACTAATCCAAAG 443
QY 564 GG-TGAGATGATTATAGTTTAAATAGTTTAAACAGGAGGAAAAA-GCTGTCTGACAG 621

DB 444 GGCTGAGATGATTAAATAGTTTAAAAAATTAGTTAAACAGAGGAAAAAGCGTCCCTGACAG 503
QY 622 CCAGTCAAGTATCTTTTACCTGTGTCGGAATGATTCGTGTCGTGTCGATTTTAAATATT 681
DB 504 CCAGTCAAGTATCTTTTACCTGTGTCGGAATGATTCGTGTCGTGTCGATTTTAAATATT 563
QY 682 TTTTGAAGAGCGGAAAAATAAGTTCTGAAGAGATAAAACCGCCTATATAAATTCATATAT 741
DB 564 TTTTGAAGAGCGGAAAAATAAGTTCTGAAGAGATAAAACCGCCTATATAAATTCATATAT 623
QY 742 TTTCTCTCCCGCTTGAATGTCTCGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 801
DB 624 TTTCTCTCCCGCTTGAATGTCTCGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 683
QY 802 TCCGGGAGCTTCACAGAGAAAGAAAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 861
DB 684 TCCGGGAGCTTCACAGAGAAAGAAAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 743
QY 862 CAGGAGATTCATTTCTCCGTTTGAATCTTTCTCAATCTCTCTCTCTCTCTCTCTCTCTCTCT 921
DB 744 CAGGAGATTCATTTCTCCGTTTGAATCTTTCTCAATCTCTCTCTCTCTCTCTCTCTCTCT 800
QY 922 TTTCCAAGGTAAATAGAACTTTCTCGATCTACTTTATTTCTCGATCTCGATCTCTGTTTT 981
DB 801 TTTCCAAGGTAAATAGAACTTTCTCGATCTACTTTATTTCTCGATCTCGATCTCTGTTTT 860
QY 982 CTCAATTTCTTGAGATCTGGAATCTGTTTAAATTT-GGATCTCTGAACTCCACTAAATC 1040
DB 861 CTCAATTTCTTGAGATCTGGAATCTGTTTAAATTTGGGATCTGGAATCTCCACTAAATC 920
QY 1041 TTTTGGTTTACTAGATTCGATCTAAAGTTGACCGATCAGTTAGTCTGATTTAGCTTACCA 1100
DB 921 TTTTGGTTTACTAGATTCGATCTAAAGTTGACCGATCAGTTAGTCTGATTTAGCTTACCA 980
QY 1101 GAATTTGGCTTGACCTTGAGAGAGATCCATGTTTCAATGTTTACCTGGAAATGATTTGTA 1160
DB 981 GAATTTGGCTTGACCTTGAGAGAGATCCATGTTTCAATGTTTACCTGGAAATGATTTGTA 1040
QY 1161 TATGTGAATTTAAATCTGAACTGTGTAAGTTAGATTTGAATCTGAACTGTCAATGTTAG 1220
DB 1041 TATGTGAATTTAAATCTGAACTGTGTAAGTTAGATTTGAATCTGAACTGTCAATGTTAG 1100
QY 1221 ATTGAATCTGAACACTGTTTAA-GTTAGATGAAGTTTGTGTATAGATTTCTTGAACCTT 1279
DB 1101 ATTGAATCTGAACACTGTTTAAAGTTAGATGAAGTTTGTGTATAGATTTCTTGAACCTT 1160
QY 1280 AGGATTTGTAGTGTGTCAGTGTGAACAGAAAGCTATTCTGATTCAATCAGGTTTATTT 1339
DB 1161 AGGATTTGTAGTGTGTCAGTGTGAACAGAAAGCTATTCTGATTCAATCAGGTTTATTT 1220
QY 1340 GACTGTATTGAACCTTTTTTGTGTTTGCAGCT 1373
DB 1221 GACTGTATTGAACCTTTTTTGTGTTTGCAGCT 1254
RESULT 13
AAD01289
ID AAD01289 standard; DNA; 17111 BP.
XX AC AAD01289;
XX DT 12-OCT-2000 (first entry)
XX Arabidopsis transformation binary vector, pAct2af-bin.
XX Matrix Attachment Region; MAR; scaffold attachment region; arabidopsis;
KW transformation binary vector; pAct2af-bin; gene expression;
KW transgenic organism; ds.
XX Synthetic.
XX Key Location/Qualifiers
FT misc_feature 50..712

FT /*tag= a
FT /note= "Corresponds to nucleotides 14-676 of MAR dimer-2"
FT misc_feature 4039..4693
FT /*tag= b
FT /note= "Corresponds to nucleotides 1-646 of MAR dimer-1"
XX WO200032800-A1.
XX 08-JUN-2000.
XX 30-NOV-1999; 99WO-US028123.
XX 01-DEC-1998; 98US-0110437P.
XX (DOWC) DOW AGROSCIENCES LLC.
XX Van Der Geest AHM, Ainley WM, Cowen NM, Welter ME, Woosley AT;
XX WPI; 2000-412345/35.
XX An isolated DNA molecule for use as a matrix attachment region to
PT increase expression of genes introduced in transformed plants comprises a
PT 298 base pair sequence described in the specification.
XX Example 3; Page 59-64; 73pp; English.
XX The patent discloses a DNA molecule, useful as matrix attachment region
CC (MAR) or scaffold attachment region, to increase the expression of genes
CC introduced in transformed plants. MARs are located in non-transcribed
CC regions of genes and form the physical boundaries of individual DNA
CC loops. They are rich in adenosine and thymine bases and contain certain
CC conserved sequence elements and structural features. They are about 300-
CC 2000 bp in length. Increased levels of expression of DNA introduced into
CC plants can be achieved by use of MAR. They can also reduce the position
CC effect in transgenic organisms. The present DNA sequence is the
CC Arabidopsis transformation binary vector pAct2af-bin, identical to the
CC vector pAct2-bin, except that it contains a MAR dimer-2 positioned 5' to
CC the Act2 transcription initiation region and the MAR dimer-1 positioned
CC 3' to the nos 3' UTR (untranslated region). This vector is used to test
CC the two orientations of the artificial MAR dimer in Arabidopsis
XX SQ Sequence 17111 BP; 4720 A; 3952 C; 4065 G; 4371 T; 0 U; 3 Other;
Query Match 74.7%; Score 1078.8; DB 3; Length 17111;
Best Local Similarity 98.3%; Pred. No. 2.2e-250;
Matches 1154; Conservative 0; Mismatches 12; Indels 8; Gaps 6;
QY 205 GATTGAGATCGTTTGGACGAGTTGGAGTGTAGTAGTGGCCATTATTTAATGTACATCT 264
DB 713 GAATTCGGCTTGTGTTGACGAGTTGGAGTGTAGTAGTGGCCATTATTTAATGTACATCT 772
QY 265 AATCGTGAATGATG-ATATGATGAACATTGTATCTTATGTTATTAATATCCATAACAC 323
DB 773 AATCGTGAATGATGATGATGAACATTGTATCTTATGTTATTAATATCCATAACAC 832
QY 324 ATCATGAAGACACATCTTCTTCAGGGTCTGAATTAATATGATCAATCTTAATAGAAA 383
DB 833 ATCATGAAGACACATCTTCTTCAGGGTCTGAATTAATATGATCAATCTTAATAGAAA 892
QY 384 CGAATTAAATACGTTGAATTTGATGAATCTTAATGAACCAAGCCACCAAGGAGAC 443
DB 893 CGAATTAAATACGTTGAATTTGATGAATCTTAATGAACCAAGCCACCAAGGAGAC 952
QY 444 TAACTGTGCTGGATGATCTCGGTTTAAGTTAAACCACTTAAACCAAGGAGTGTCAATGA 503
DB 953 TAACTGTGCTGGATGATCTCGGTTTAAGTTAAACCACTTAAACCAAGGAGTGTCAATGA 1012
QY 504 ACACGCGGATCGGAGGAGTGTCACTCATGAAGCCATCAAGCAAAAAGAACTAATCCAAAG 563
DB 1013 ACACGCGGATCGGAGGAGTGTCACTCATGAAGCCATCAAGCAAAAAGAACTAATCCAAAG 1072
QY 564 GG-TGAGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 621
|||

DB 1073 GCGTGAGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1132
QY 622 CCAGGTCACTTATCTTTACCTGTGTGCGAATGATTCGTGTCTGTCTGATTTTAAATTAAT 681
DB 1133 CCAGGTCACTTATCTTTACCTGTGTGCGAATGATTCGTGTCTGTCTGATTTTAAATTAAT 1192
QY 682 TTTTGTAAAGCGCCGAAATAAAGTTGTAAGAGATAAACCCTCTATATAAATTCATATAT 741
DB 1193 TTTTGTAAAGCGCCGAAATAAAGTTGTAAGAGATAAACCCTCTATATAAATTCATATAT 1252
QY 742 TTTTCTCCCGCTTTTGAATTTGTCTGTGTCTCTCACTTTTCAATCAGCCGTTTGAATC 801
DB 1253 TTTTCTCTCCGCTTTTGAATTTGTCTGTGTCTCTCACTTTTCAATCAGCCGTTTGAATC 1312
QY 802 TCCGGGCACTTGCAGAGNAGAACAGGAGAGAGTAAGAGAGAAAGTAAGAGATAATC 861
DB 1313 TCCGGGCACTTGCAGAGAGAAACAAGAGAGAGAAAGTAAGAGAGATAATC 1372
QY 862 CAGGAGATTCATTTCTCCGCTTTTGAATTTGTCTCTCAATCTCATCTTCTTCTCCGCTCTTTT 921
DB 1373 CAGGAGATTCATTTCTCCGCTTTTGAATTTGTCTCTCAATCTCA---TCTTCTCCGCTCTTTT 1429
QY 922 TTTTCAAGGTAATAGGAATCTTTCTGGAATCTACTTTATTTTGTGATCTCGATCTTGTGTTT 981
DB 1430 TTTTCAAGGTAATAGGAATCTTTCTGGAATCTACTTTATTTTGTGATCTCGATCTTGTGTTT 1489
QY 982 CTCATATTTCTTGAGATCTGGAATTCGTTTAAATTT---GGATCTGTGAACCTCCACTAAATC 1040
DB 1490 CTCATATTTCTTGAGATCTGGAATTCGTTTAAATTTTGGGATCTGTGAACCTCCACTAAATC 1549
QY 1041 TTTTGGTTTACTAGAAATCGATCTTAAGTTGACCGATCAGTTAGCTCGATTTATAGTACCA 1100
DB 1550 TTTTGGTTTACTAGAAATCGATCTTAAGTTGACCGATCAGTTAGCTCGATTTATAGTACCA 1609
QY 1101 GAATTTGGCTTGACCTTGATGAGAGATCCATGTTTCATGTTTACCTGGAAATGATTGTA 1160
DB 1610 GAATTTGGCTTGACCTTGATGAGAGATCCATGTTTCATGTTTACCTGGAAATGATTGTA 1669
QY 1161 TATGTGAATTTGAATCTGAATCTGTTTGAAGTTAGATTGAATCTGAACACTGTCAATGTTAG 1220
DB 1670 TATGTGAATTTGAATCTGAATCTGTTTGAAGTTAGATTGAATCTGAACACTGTCAATGTTAG 1729
QY 1221 ATTGAATCTGAACACTGTGTTTAA---GTTAGATGAAGTTTGTGTATAGATTCTTCCAAACCTT 1279
DB 1730 ATTGAATCTGAACACTGTGTTTAAAGTTAGATTGAAGTTTGTGTATAGATTCTTCCAAACCTT 1789
QY 1280 AGGATTTGTAGTGTGTAGCTGTGAACAGAGATTTTCTGATTCATCAATCAGGGTTTATTT 1339
DB 1790 AGGATTTGTAGTGTGTAGCTGTGAACAGAGATTTTCTGATTCATCAATCAGGGTTTATTT 1849
QY 1340 GACTGTATTGAACCTCTTTTGTGTTTGCAGCT 1373
DB 1850 GACTGTATTGAACCTCTTTTGTGTTTGCAGCT 1883
RESULT 14
AAD01290
ID AAD01290 standard; DNA; 17116 BP.
XX
AC AAD01290;
XX
DT 12-OCT-2000 (first entry)
XX
DE Arabidopsis transformation binary vector, pAct2af-bin.
XX
KW Matrix Attachment Region; MAR; scaffold attachment region; pAct2af-bin;
KW Arabidopsis transformation binary vector; gene expression;
XX transgenic organism; ds.
XX
OS Synthetic.
XX
XX Key Location/Qualifiers
FT misc_feature 64..709

1078	DB	GCCTCAGATGATTTAAATTAGTTTAAATAATTAGTTAAACAGAGGGAAGGGCTGCCTGCACAG	1133
622	QY	CCAGGTACAGGTATCTTTTACCTCGTGGTGGAAATGATTCGTGTCGTGATTTTAAATATT	681
1138	DB	CCAGGTACAGGTATCTTTTACCTCGTGGTGGAAATGATTCGTGTCGTGATTTTAAATATT	1197
682	QY	TTTTTTGAAGGCCGAAATAAAGTCTGAAGAGATAAAACCGGCTATATAAATTCATATAT	741
1198	DB	TTTTTTGAAGGCCGAAATAAAGTCTGAAGAGATAAAACCGGCTATATAAATTCATATAT	1257
742	QY	TTTTCTCCCGCTTTGAATTCGTCTCGTGTCTCTCACTTTTCATCAGCCGTTTGAATC	801
1258	DB	TTTTCTCTCCCGCTTTGAATTCGTCTCGTGTCTCTCACTTTTCATCAGCCGTTTGAATC	1317
802	QY	TCGCGGCACTTGACAGAGAACAACAGGAAGAAGACATTAAGAGAGAAAGTAAGAGATAATC	861
1318	DB	TCGCGGCACTTGACAGAGAACAACAGGAAGAAGACATTAAGAGAGAAAGTAAGAGATAATC	1377
862	QY	CAGGAGATTCATTTCTCCGTTTTGAATCTTCTCCAAATCTCATCTTCTTCGCGCTCTTTC	921
1378	DB	CAGGAGATTCATTTCTCCGTTTTGAATCTTCTCCAAATCTCA--TCTCTTCGCGCTCTTTC	1434
922	QY	TTTTCAAAGTAAATAGCACTTCTCGAATCTACTTTATTTCTCTGGATCTCGATCTCGTTTT	981
1435	DB	TTTTCAAAGTAAATAGCACTTCTCGAATCTACTTTATTTCTCTGGATCTCGAATCTGTTTT	1494
982	QY	CTCAATTTTCTCTCAGATCTGGAAATCTGGTTTAAATTT--GGATCTGTGAAACCTCCATAAATC	1040
1495	DB	CTCAATTTTCTCTCAGATCTGGAAATCTGGTTTAAATTTGGGATCTGTGAACCTCCATAAATC	1554
1041	QY	TTTTGGTTTTATAGAAATCGATCTAAGTTGACCGATCAGTTAGCTCGAATATAGCTACCA	1100
1555	DB	TTTTGGTTTTATAGAAATCGATCTAAGTTGACCGATCAGTTAGCTCGAATATAGCTACCA	1614
1101	QY	GAATTTGCGCTTGACCTTGATGGAGAGATCCATGTTTCATGTTACCTGGGAAATGATTTGTA	1160
1615	DB	GAATTTGCGCTTGACCTTGATGGAGAGATCCATGTTTCATGTTACCTGGGAAATGATTTGTA	1674
1161	QY	TATGTGAATTGAATCTGAACCTGTTGAAGTTAGATTTGAATCTGGAACACGTGCAATGTTAG	1220
1675	DB	TATGTGAATTGAATCTGAACCTGTTGAAGTTAGATTTGAATCTGGAACACGTGCAATGTTAG	1734
1221	QY	ATTGAATCTGAACACCTGTTTAA--CTTAGATGAAGTTTGCTATAGATTTCTTCGAAACCTT	1279
1735	DB	ATTGAATCTGAACACCTGTTTAAAGGTTAGATGAAGTTTGCTATAGATTTCTTCGAAACCTT	1794
1280	QY	AGGATTTGATGTCTGATCTGATTTGAACAGAAAGCTATTTCTGATTCAAATCAGGGTTTATT	1339
1795	DB	AGGATTTGATGTCTGATCTGATTTGAACAGAAAGCTATTTCTGATTCAAATCAGGGTTTATT	1854
1340	QY	GACTGTATGAACTCTTTTGTGTGTTTCGACCT	1373
1855	DB	GACTGTATGAACTCTTTTGTGTGTTTCGACCT	1888
RESULT 15			
ABK52077			
ID	ABK52077	standard; DNA; 1202 BP.	
XX	AC		
XX	AC	ABK52077;	
DT	DT	13-AUG-2002 (first entry)	
XX	DE	Modified plant promoter #2.	
XX	KW	Promoter; ss; plant; soybean; peanut; alfalfa; tomato; eggplant; potato;	
XX	KW	cabbage; turnip; rapeseed; apple; pear; berry; cucumber; carrots;	
XX	KW	transgenic plant.	
XX	OS	Synthetic.	
XX	PN	US2002049992-A1.	
XX	PN		

Result No.	Score	Query #	Match	Length	DB	ID	Description
1	1174.8	81.4	1219	4	US-09-737-698B-9	Sequence 9, Appl	
2	1174.8	81.4	1219	4	US-09-737-626A-9	Sequence 9, Appl	
3	1174.8	81.3	1742	4	US-09-737-698B-30	Sequence 30, Appl	
4	1174	81.3	1742	4	US-09-737-626A-30	Sequence 30, Appl	
5	75.4	5.2	1271	4	US-09-737-698B-10	Sequence 10, Appl	
6	75.4	5.2	1271	4	US-09-737-626A-10	Sequence 10, Appl	
7	73.8	5.1	1800	4	US-09-737-698B-29	Sequence 29, Appl	
8	73.8	5.1	1800	4	US-09-737-626A-29	Sequence 29, Appl	
9	50.4	3.5	6070	4	US-10-204-708-4	Sequence 9, Appl	
10	48.6	3.4	731	1	US-08-451-405A-2	Sequence 2, Appl	
11	47.9	3.3	5203	4	US-09-251-770-1	Sequence 1, Appl	
12	46.6	3.2	640681	4	US-09-790-988-1	Sequence 1, Appl	
13	46.2	3.2	2448	1	US-08-526-964-2	Sequence 2, Appl	
14	46.2	3.2	2448	2	US-08-946-617-2	Sequence 2, Appl	
15	46.2	3.2	2448	3	US-09-031-897-2	Sequence 2, Appl	
16	45.4	3.1	7218	1	US-08-232-463-14	Sequence 14, Appl	
17	45.2	3.1	1705	1	US-08-090-533-25	Sequence 25, Appl	
18	45.2	3.1	1705	1	US-08-334-639-4	Sequence 4, Appl	
19	45.2	3.1	1705	1	US-08-398-637-25	Sequence 25, Appl	
20	45.2	3.1	1705	1	US-08-406-858-26	Sequence 26, Appl	
21	45.2	3.1	1705	4	US-08-399-023-25	Sequence 25, Appl	
22	45.2	3.1	1705	5	PCT-US94-05275-26	Sequence 26, Appl	
23	44.4	3.1	7218	1	US-08-232-463-14	Sequence 14, Appl	
24	43.4	3.0	287	4	US-09-313-294A-6702	Sequence 6702, Ap	
25	42.8	3.0	274	4	US-09-313-294A-2611	Sequence 2611, Ap	
26	42.6	3.0	774	4	US-09-314-001C-725	Sequence 725, App	
27	42.4	2.9	288	4	US-09-313-294A-4386	Sequence 4386, Ap	

301	CGTTTTAAGTTAACACATAAAAAACGGAGCTGTCATGTAAACACGGGATTCGAGCGGTC	360
524	ACAGTCATGAAGCCATCAAAGCAAAAGAACTAATCAAGGGGTGAGATGATTAATTAGTT	583
361	ACAGTCATGAAGCCATCAAAGCAAAAGAACTAATCAAGGGGTGAGATGATTAATTAGTT	420
584	TAAAAATTAGTTAAACAGAGGGAAAAA-GCTGCTGACACCGCAGGTCAAGTTACTTTACC	642
421	TAAAAATTAGTTAAACAGAGGGAAAAAGGCTGCTCTGCACGACAGGTCAAGTTACTTTACC	480
643	TCGTGTCGAAATGATTTCTGTGTCGATTTTAAATATATTTTGAAGGCGGAAAAATAA	702
481	TGTGTCGAAATGATTTCTGTGTCGATTTTAAATATATTTTGAAGGCGGAAAAATAA	540
703	AGTTGTGAAGAGATAAACCCGCTATATAAATTCATATATTTTCTCCCGCTTTGAATG	762
541	AGTTGTGAAGAGATAAACCCGCTATATAAATTCATATATTTTCTCTCCGCTTTGAATG	600
763	TCTCGTTGTCCTCCTCACTTTCATCAGCGTTTGAATCTCCGCGCACTTGACAGAGAG	822
601	TCTCGTTGTCCTCCTCACTTTCATCAGCGTTTGAATCTCCGCGCACTTGACAGAGAG	660
823	AACAGGAAAGAGACTTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATCTCCGTTT	882
661	AACAGGAAAGAGACTTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCATCTCCGTTT	720
883	TGAATCTCCCTCAATCTCATCTTCTCTCTCCGCTCTTTCTTTCCAGGTAATAGGAATCT	942
721	TGAATCTCCCTCAATCTCA--TCTCTCTCCGCTCTTTCTTTCCAGGTAATAGGAATCT	777
943	TCCTGAGTACTATTTATTTGCTGAGATCTCGATCTGTTTCTCAATTTCTTCTGAGATCTG	1002
778	TCCTGAGTACTATTTATTTGCTGAGATCTCGATCTGTTTCTCAATTTCTTCTGAGATCTG	837
1003	AATTCGTTTAAATTGGATCTGTGAACCTCCACTAAATCTTTTGGTTTACTAGAACTCGAT	1062
838	AATTCGTTTAAATTGGATCTGTGAACCTCCACTAAATCTTTTGGTTTACTAGAACTCGAT	897
1063	CTAAGTTGACCGATCAGTTAGCTCGATTAAGCTACACAGAAATTTGGCTGACCTTGATGG	1122
898	CTAAGTTGACCGATCAGTTAGCTCGATTAAGCTACACAGAAATTTGGCTGACCTTGATGG	957
1123	AGAGATCCATGTTCAATGTTTACTGGGAATGATTTGTATGTGAATTTGAAATCTGAACT	1182
958	AGAGATCCATGTTCAATGTTTACTGGGAATGATTTGTATGTGAATTTGAAATCTGAACT	1017
1183	GTTGAAGTTAGATTGAATCTGAACCTGCAACTGTCAATGTGTAGATTCGAATCTGAACT	1242
1018	GTTGAAGTTAGATTGAATCTGAACCTGCAACTGTCAATGTGTAGATTCGAATCTGAACT	1077
1243	GTTAGATCAAGTTTGTATAGATTTCTTGGAACTTTAGATTTGATGTCGTCAGTTG	1302
1078	GTTAGATCAAGTTTGTATAGATTTCTTGGAACTTTAGATTTGATGTCGTCAGTTG	1137
1303	AACAGAAAGCTATTCTGATTCATCAGGGTTTATTTGACTGTATTGAATCTTTTTTGTG	1362
1138	AACAGAAAGCTATTCTGATTCATCAGGGTTTATTTGACTGTATTGAATCTTTTTTGTG	1197
1363	TGTTTGCAGCTCATAAAAATG	1384
1198	TGTTTGCAGAGACTCACCATG	1219

RESIN, T. 2

US-09-737-626A-9
; Sequence 9, Application US/09737626A
; Patent No. 6660911
; GENERAL INFORMATION:
; APPLICANT: Fincher, Karen
; APPLICANT: Flasiński, Stanisław
; APPLICANT: Wilkinson, Jack
; TITLE OF INVENTION: No. 6660911el p

RESULT 4

US-09-737-626A-30
; Sequence 30, Application US/09737626A
; Patent No. 6660911
; GENERAL INFORMATION:
; APPLICANT: Fincher, Karen
; APPLICANT: Flasiński, Stanisław
; APPLICANT: Wilkinson, Jack
; TITLE OF INVENTION: No. 6660911el Plant Expression Constructs
; FILE REFERENCE: 38-21(51499)B
; CURRENT APPLICATION NUMBER: US/09/737,626A
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 09/737,626
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 30
; LENGTH: 1742
; TYPE: DNA
; ORGANISM: artificial sequence
; NAME/KEY: promoter
; LOCATION: (1)..(1742)
; OTHER INFORMATION: chimeric promoter fusion CamV and Act2 polynucleotides + Act2 in
; OTHER INFORMATION: tro
US-09-737-626A-30

Query Match 81.3%; Score 1174; DB 4; Length 1742;
Best Local Similarity 99.3%; Pred. No. 1.5e-305;
Matches 1201; Conservative 0; Mismatches 5; Indels 4; Gaps 2;
QY 164 CAACATATTTTATGATGACGAGTTCAGCATATGATTAATGATTCAGAACTGTTTAC 223
DB 533 CAACATATTTTATGATGACGAGTTCAGCATATGATTAATGATTCAGAACTGTTTAC 592
QY 224 GAGTTCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 283
DB 593 GAGTTCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 652
QY 284 ATGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 343
DB 653 ATGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 712
QY 344 TCAGGCTGTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 403
DB 713 TCAGGCTGTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 772
QY 404 TGATGAAATCTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 463
DB 773 TGATGAAATCTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 832
QY 464 CGTTTAAGTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 523
DB 833 CGTTTAAGTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 892
QY 524 ACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 583
DB 893 ACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 952
QY 584 TAAATAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 642
DB 953 TAAATAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1012
QY 643 TGTGCTGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 702
DB 1013 TGTGCTGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1072
QY 703 AGTTGTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 762
DB 1073 AGTTGTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1132

QY 763 TCTGTTGCTCTCTCACTTTTCATCAGCGTGTGTTTGAATCTCCGGGACCTTCACAGAGA 822
DB 1133 TCTGTTGCTCTCTCACTTTTCATCAGCGTGTGTTTGAATCTCCGGGACCTTCACAGAGA 1192
QY 823 AACAGGAAGAAGACTAAG 882
DB 1193 AACAGGAAGAAGACTAAG 1252
QY 883 TGAATCTTCTCAATCTCAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 942
DB 1253 TGAATCTTCTCAATCTCAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1309
QY 943 TCTGGATCTACTTATTTTGTCTGGATCTCGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1002
DB 1310 TCTGGATCTACTTATTTTGTCTGGATCTCGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1369
QY 1003 AATTCTGTTTAAATTTGGATCTGTGAACCTCCACTAAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1062
DB 1370 AATTCTGTTTAAATTTGGATCTGTGAACCTCCACTAAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1429
QY 1063 CTAAGTTGACCGATCAGTTAGCTGATTAATAGCTACAGAAATTTGGCTTGACCTTGATGG 1122
DB 1430 CTAAGTTGACCGATCAGTTAGCTGATTAATAGCTACAGAAATTTGGCTTGACCTTGATGG 1489
QY 1123 AGAGATCCATGTTTCAATGTTTACCTGGGAATGATTTGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1182
DB 1490 AGAGATCCATGTTTCAATGTTTACCTGGGAATGATTTGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1549
QY 1183 GTTGAAGTTAGTAAATCTGATG 1242
DB 1550 GTTGAAGTTAGTAAATCTGATG 1609
QY 1243 GTTGAAGTTAGTAAATCTGATG 1302
DB 1610 GTTGAAGTTAGTAAATCTGATG 1669
QY 1303 AACAGGAAGCTATTTCTGATG 1362
DB 1670 AACAGGAAGCTATTTCTGATG 1729
QY 1363 TGTTTGCAGC 1372
DB 1730 TGTTTGCAGC 1739

RESULT 5

US-09-737-698B-10
; Sequence 10, Application US/09737698B
; Patent No. 6462258
; GENERAL INFORMATION:
; APPLICANT: Fincher, Karen
; APPLICANT: Wilkinson, Jack
; TITLE OF INVENTION: No. 6462258el Plant Expression Constructs
; FILE REFERENCE: 38-21(51499)C
; CURRENT APPLICATION NUMBER: US/09/737,698B
; CURRENT FILING DATE: 2000-12-15
; PRIOR APPLICATION NUMBER: US 60/171,173
; PRIOR FILING DATE: 1999-12-16
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 1271
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; NAME/KEY: promoter
; LOCATION: (1)..(1271)
; OTHER INFORMATION: y = t/u or c
; OTHER INFORMATION: Act8 promoter polynucleotide sequence and intron
US-09-737-698B-10

Query Match 5.2%; Score 75.4; DB 4; Length 1271;
Best Local Similarity 51.6%; Pred. No. 1.2e-10;

Matches 214; Conservative 3; Mismatches 194; Indels 4; Gaps 2;

QY 606 AAAAGCTGTCTGACAGCCAGGTCAACGTTA---TCTTTACCTGTGGTGCAGAAATGATTCGTG 662
DB 471 AGAGGCTGTCTGACAGCCATGTCACTTTATCTTTTCCGTATGATCGAAATGATTCGTG 530
QY 663 TCTGTGATTTTAAATTTATTTTGAAGCCGGAATAAAGTTGTAAGAGATAAACC 722
DB 531 TTTGAGATTTTAAATTTATTTTCCAAATTAAGACTCTAAGAAAAAATAGTTTTC 590
QY 723 CCTATATAAATTCATATATTTTCTCCCGCTTTGAATGTCGTGTTCTCTCACTT 782
DB 591 GATAAACCCGCTATATAAATAGTCAACACTCGGTTTATTTCTCTCCCTCAAGAAAT 650
QY 783 TCATCAGCGTTTGAATCTCCGCGACTTGACAGAGAAACAGGAGAGACTAAGA 842
DB 651 TGCCTGCTGCTTCAAGTTCATCGCTCGTGTCTTCTTCCTCGATCTCTCGA 770
QY 843 -GAGAAAGTAAGAGATAATCCAGAGATTCATCTCCGTTTGAATCTCTCAATCTCA 901
DB 711 GGAGAAAGAGTGAGCCAGTCTTCTATCGCTCGTGTCTTCTTCCTCGATCTCTCGA 770
QY 902 TCTTCTTCTCCGCTTCTTCTTCCAAAGTAATAGGAACCTTCTCGATCTACTTTATTG 961
DB 771 TCTTCTGCTTGTCTTCCGATTAAGGTAATTAAGAACTCCGATCTACTTGTCTGTG 830
QY 962 CTGATCTGATCTTGTCTTCTCAATTTCTGAGATCTGGAATTCGTTTAAATTT 1016
DB 831 TTGATCTGATCTGATTTCTAAGTTTACCTTCAAAAGTTGTTTCCGATTTGATT 885

RESULT 6

US-09-737-626A-10

; Sequence 10, Application US/09737626A

; Patent No. 6660911

; GENERAL INFORMATION:

; APPLICANT: Flinski, Karen

; APPLICANT: Wilkins, Jack

; TITLE OF INVENTION: No. 6660911 Plant Expression Constructs

; FILE REFERENCE: 38-21(51499)B

; CURRENT APPLICATION NUMBER: US/09/737,626A

; PRIOR FILING DATE: 2002-02-25

; PRIOR APPLICATION NUMBER: 09/737,626

; PRIOR FILING DATE: 2000-12-15

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 10

; LENGTH: 1271

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

; FEATURE:

; NAME/KEY: promoter

; LOCATION: (1)..(1271)

; OTHER INFORMATION: y = t/u or c

; OTHER INFORMATION: Act8 promoter polynucleotide sequence and intron

US-09-737-626A-10

Query Match 5.2%; Score 75.4; DB 4; Length 1271;

Best Local Similarity 51.6%; Pred. No. 1.2e-10;

Matches 214; Conservative 3; Mismatches 194; Indels 4; Gaps 2;

QY 606 AAAAGCTGTCTGACAGCCAGGTCAACGTTA---TCTTTACCTGTGGTGCAGAAATGATTCGTG 662
DB 471 AGAGGCTGTCTGACAGCCATGTCACTTTATCTTTTCCGTATGATCGAAATGATTCGTG 530
QY 663 TCTGTGATTTTAAATTTATTTTGAAGCCGGAATAAAGTTGTAAGAGATAAACC 722
DB 531 TTTGAGATTTTAAATTTATTTTCCAAATTAAGACTCTAAGAAAAAATAGTTTTC 590
QY 723 CCTATATAAATTCATATATTTTCTCCCGCTTTGAATGTCGTGTTCTCTCACTT 782
DB 591 GATAAACCCGCTATATAAATAGTCAACACTCGGTTTATTTCTCTCCCTCAAGAAAT 650

QY 783 TCATCAGCGTTTGAATCTCCGCGACTTGACAGAGAAACAGGAGAGACTAAGA 842
DB 651 TGCCTGCTGCTTCAAGTTCATCGCTCGTGTCTTCTTCCTCGATCTCTCGA 770
QY 843 -GAGAAAGTAAGAGATAATCCAGAGATTCATCTCCGTTTGAATCTCTCAATCTCA 901
DB 711 GGAGAAAGAGTGAGCCAGTCTTCTATCGCTCGTGTCTTCTTCCTCGATCTCTCGA 770
QY 902 TCTTCTTCTCCGCTTCTTCTTCCAAAGTAATAGGAACCTTCTCGATCTACTTTATTG 961
DB 771 TCTTCTGCTTGTCTTCCGATTAAGGTAATTAAGAACTCCGATCTACTTGTCTGTG 830
QY 962 CTGATCTGATCTTGTCTTCTCAATTTCTGAGATCTGGAATTCGTTTAAATTT 1016
DB 831 TTGATCTGATCTGATTTCTAAGTTTACCTTCAAAAGTTGTTTCCGATTTGATT 885

RESULT 7

US-09-737-698B-29

; Sequence 29, Application US/09737698B

; Patent No. 6462258

; GENERAL INFORMATION:

; APPLICANT: Fincher, Karen

; APPLICANT: Wilkins, Jack

; TITLE OF INVENTION: No. 6462258 Plant Expression Constructs

; FILE REFERENCE: 38-21(51499)C

; CURRENT APPLICATION NUMBER: US/09/737,698B

; PRIOR FILING DATE: 2000-12-15

; PRIOR APPLICATION NUMBER: US 60/171,173

; PRIOR FILING DATE: 1999-12-16

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 29

; LENGTH: 1800

; TYPE: DNA

; ORGANISM: artificial sequence

; FEATURE:

; NAME/KEY: promoter

; LOCATION: (1)..(1800)

; OTHER INFORMATION: y = t/u or c

; OTHER INFORMATION: chimeric promoter fusion CamV and Act8 polynucleotides + Act8 ir

US-09-737-698B-29

Query Match 5.1%; Score 73.8; DB 4; Length 1800;

Best Local Similarity 51.3%; Pred. No. 3.7e-10;

Matches 213; Conservative 3; Mismatches 195; Indels 4; Gaps 2;

QY 606 AAAAGCTGTCTGACAGCCAGGTCAACGTTA---TCTTTACCTGTGGTGCAGAAATGATTCGTG 662
DB 1004 AGAGGCTGTCTGACAGCCATGTCACTTTTCCGTATGATCGAAATGATTCGTG 1063
QY 663 TCTGTGATTTTAAATTTATTTTGAAGCCGGAATAAAGTTGTAAGAGATAAACC 722
DB 1064 TTTGAGATTTTAAATTTATTTTCCAAATTAAGACTCTAAGAAAAAATAGTTTTC 1123
QY 723 CCTATATAAATTCATATATTTTCTCCCGCTTTGAATGTCGTGTTCTCTCACTT 782
DB 1124 GATAAACCCGCTATATAAATAGTTCACACTCGGTTTATTTCTTCTCCCTCAAGAAAT 1183
QY 783 TCATCAGCGTTTGAATCTCCGCGACTTGACAGAGAAACAGGAGAGACTAAGA 842
DB 1184 TGCCTGCTGCTTCAAGTTCATCGCTCGTGTCTTCTTCCTCGATCTCTCGA 1243
QY 843 -GAGAAAGTAAGAGATAATCCAGAGATTCATCTCCGTTTGAATCTCTCAATCTCA 901
DB 1244 GGAGAAAGAGTGAGCCAGATCTTCACTCGTGTGTTTCTTCTTCCTCGATCTCTCGA 1303
QY 902 TCTTCTTCTCCGCTTCTTCTTCCAAAGTAATAGGAACCTTCTCGATCTACTTTATTG 961
DB 1304 TCTTCTGCTTGTCTTCCGATTAAGTTTAAAGCTCCGATCTACTTGTCTGTG 1363

QY 962 CTGGATCTCGATCTGTTCTTCAATTTCTTGGAGATCTGGAATCTGTTAAATTT 1016
|||||
Db 1364 TTGGATCTCGATCTGATTAAGTTTACCTTCAAAAGTTGTTCCGATTTGATT 1418
|||||

RESULT 8

US-09-737-626A-29
; Sequence 29, Application US/09737626A
; Patent No. 6660911
; GENERAL INFORMATION:
; APPLICANT: Fincher, Karen
; APPLICANT: Flasiński, Stanisław
; APPLICANT: Wilkinson, Jack
; TITLE OF INVENTION: No. 6660911el Plant Expression Constructs
; FILE REFERENCE: 38-21(51499)B
; CURRENT APPLICATION NUMBER: US/09/737,626A
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 09/737,626
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 29
; LENGTH: 1800
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1800)
; OTHER INFORMATION: Y = t/u or c
; OTHER INFORMATION: chimeric promoter fusion CamV and Act8 polynucleotides + Act8 in
; OTHER INFORMATION: ttc
US-09-737-626A-29

Query Match 5.1%; Score 73.8; DB 4; Length 1800;
Best Local Similarity 51.3%; Pred. No. 3.7e-10;
Matches 213; Conservative 3; Mismatches 195; Indels 4; Gaps 2;

QY 606 AAGAGCTGCTGACAGCCAGTCACTTA---TCTTTTACCTGGTTCGAATGATTCGG 662
|||||
Db 1004 AGAGGCTGCTGACAGCCAGTCACTTA---TCTTTTACCTGGTTCGAATGATTCGG 1063
|||||

QY 663 TCTGTCGATTTAAATTTTAAAGGCGGAAAATAAGTTGTAAGAGATAAAACCG 722
|||||
Db 1064 TTGGYGAATTTAAATTTTCCAAATTCAGTCACTTAAGAAAATAAGTTTTC 1123
|||||

QY 723 CTTATATAAATCAATATTTCTCCCGCTTTGAAATGCTCTCGTCTCTCTCACTT 782
|||||
Db 1124 GATAAACCGGCTATATAAATAGTTCAACACTCGGTTTATTTCTTCCCTCRAAGAA 1183
|||||

QY 763 TCATCAGCGGTTTGAATCTCCGCGACTTGACAGAGAAGCAAGGAAGAACTAAGA 842
|||||
Db 1184 TGCTCTGCTCTTCAGCTTCATCGCGCTTGCAATTTCCCGGATTAAGAGAGAAAG 1243
|||||

QY 843 -GAGAAAGTAAGAGATAATCCAGAGATTCATCTCCGTTTGAATCTCTCAATCTCA 901
|||||
Db 1244 GGAGAAAGAGTGAGCCAGATCTTCATCGCTGGTGTCTGTTCTCTCTCGATCTCTCGA 1303
|||||

QY 902 TCTTCTCTTCGCTCTTTCTTTCGAGGTAAAGTAACTTCTCGATCTACTTTATTG 961
|||||
Db 1304 TCTTCTGCTTTTGTCTTCCGATTAAGTAAATTAACCTCCGATCTACTTGTCTGTG 1363
|||||

QY 962 CTGGATCTCGATCTGTTTCTCAATTTCTTGGAGATCTGGAATCTGTTAAATTT 1016
|||||
Db 1364 TTGGATCTCGATCTGATTAAGTTTACCTTCAAAAGTTGTTCCGATTTGATT 1418
|||||

RESULT 9

US-10-204-708-9
; Sequence 9, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian

; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: By Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 9
; LENGTH: 6070
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-9

Query Match 3.5%; Score 50.4; DB 4; Length 6070;
Best Local Similarity 44.9%; Pred. No. 0.0012;
Matches 192; Conservative 0; Mismatches 236; Indels 0; Gaps 0;

QY 838 TAAGAGAGAAAGTAAGAGATAAATCCAGAGATCAATCTCCGTTTGAATCTTCCCTCAAT 897
|||||
Db 3589 TAGTAGAGATTCATGTTTAAAGTGGTTGATTATTTATATATTTTATTAATATGCTAT 3648
|||||

QY 898 CTGATCTCTTCTCCGCTCTTCTTCCAGGTAATAGCAACTTCTCGATCTACTTTA 957
|||||
Db 3649 GAGAGTTTATTTTGTATTTTGTAAATTAAGGGAATGTTAGTTTATTTTATTTAGTTA 3768
|||||

QY 958 TTTCTCGATCTCGATCTGTTTCTCAATTTCTTCCGATCTGGAATTCGTTTAAATTTG 1017
|||||
Db 3709 TTTTATGAGTGTAAATAGTATTTATTTTGGTTTAAATTTATTTATAGAGATTTTATG 3768
|||||

QY 1018 GATCTGTGAACCTCCATTAATCTTTTGGTTTACTAGATCGATCAAGTTGACCGATC 1077
|||||
Db 3769 GAGGAGGTAATTCGAATAGATTTTATATTTTAAAGAAATTTTTCGCTATTTATGCTG 3828
|||||

QY 1078 AGTTAGCTCGATATAGTACCCGAAATTTGCTTGACCTTGATGGAGAGATCCCATGTCA 1137
|||||
Db 3829 AGAATGATTTATTTGATTAATTAAGTTTATAGTAGTATTTATTAATAGTTAAAGGTG 3888
|||||

QY 1138 TGTTCACCTGGAAATGATTTGTATATGTAATCTGAAATCTGAACTGTTGAAAGTTAGATTG 1197
|||||
Db 3889 GAATAATTAAGTATTTATTTATATGGAATGAATGGAATAAATAAATAGTATATATGAT 3948
|||||

QY 1198 AATCTGACACTGTCATAGTTAGATTCGAACACTGTTTAAAGTTAGTAGAAGTTTG 1257
|||||
Db 3949 AATGGAATATTTATTTAGTTTAAAGAGGAAGAAATTTTGATATATATGTTATATATAGATG 4008
|||||

QY 1258 TGTATAGA 1265
Db 4009 TTTTTCGA 4016

RESULT 10

US-08-451-405A-2/c
; Sequence 2, Application US/08451405A
; Patent No. 5736358
; GENERAL INFORMATION:
; APPLICANT: FASEL, NICOLAS JOSEPH
; APPLICANT: REYMOND, CHRISTOPHE DOMINIQUE
; TITLE OF INVENTION: DICTYOSTELID EXPRESSION VECTOR AND
; TITLE OF INVENTION: METHOD FOR EXPRESSING A DESIRED PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:

ADDRESSEE: THE WEBB LAW FIRM
STREET: 700 KOPPERS BUILDING, 436 SEVENTH AVENUE
CITY: PITTSBURGH

STATE: PENNSYLVANIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 15219-1818

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISK

COMPUTER: Midwest Micro 486-50
OPERATING SYSTEM: DOS
SOFTWARE: WORDPERFECT 6.1

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/451,405A

FILING DATE: 26-MAY-1995
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/965,273

FILING DATE: 15-JAN-1993
INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 731

TYPE: NUCLEIC ACID
STRANDEDNESS: SINGLE

TOPOLOGY: UNKNOWN
US-08-451-405A-2

Query Match 3.4%; Score 48.6; DB 1; Length 731;
Best Local Similarity 49.0%; Pred. No. 0.0014;
Matches 129; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

168 TATTTTATGATGCAAGAGTCAGCATATGTAATTAATTCAGATCGTTTCGACGAGT 227
Db 698 TTTTATTAAGAAAAAATAAATTTTGAATGATTAAAGAAAAAATAAATAAATAAT 639

228 TCGAGTAGTAGTAGGACATATTAATGTCATCACTAATCGTGAATAGTAGTATGATGA 287
Db 638 AATAATGTGAAAAAGGATTTTATTAATAAAGAAATTAATTAATTAATTAATTAAT 579

288 AACATCTGATCTATTGTATATAATCCATATAACACATCATGAAGACACTTTCTTCAG 347
Db 578 TATTTTATTTCTAATGATATATATAAATAAATAAATAAATAAATAAATAAATAA 519

348 GGTCTGAATTAATTAATGATACAAATCTTAATAGAAAAAGAAATTAATTAATTAAT 407
Db 518 TAAAGGGTGTGCTGTAAGAAAAAATAAATAAATAAATAAATAAATAAATAAATAA 459

408 TGAATCTAATTAATGAACCAAA 430
Db 458 AAAATAATAATTAATAAATAAATAA 436

RESULT 11
US-09-257-770-1/c

Sequence 1, Application US/09257770
Patent No. 6306596

GENERAL INFORMATION:
APPLICANT: Lambowitz, Allen M.

APPLICANT: Zimmerly, Steven
APPLICANT: Guo, Huatuo

APPLICANT: Mohr, Georg
APPLICANT: Beall, Clifford J.

TITLE OF INVENTION: Methods for Cleaving Single-Stranded and
Double-Stranded DNA Substrates with Nucleotide

TITLE OF INVENTION: Integrase
FILE REFERENCE: 24671/04007

CURRENT APPLICATION NUMBER: US/09/257,770
CURRENT FILING DATE: 1999-02-25

NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 1
LENGTH: 5203

TYPE: DNA
ORGANISM: S. cerevisiae

US-09-257-770-1

Query Match 3.3%; Score 47.8; DB 4; Length 5203;
Best Local Similarity 45.8%; Pred. No. 0.0056;
Matches 160; Conservative 0; Mismatches 189; Indels 0; Gaps 0;

73 AATATGNNACGAAAAATGGCTACACTCGATGTAATGGTATCTCAACTCAACATTATAC 132
Db 2378 AATATACCTCTATGTAATTTGTTTAAACATGATGTAATTTCTACATCAATAGTGAATTA 2319

133 TTATACCAACAACTTAGTAGCAAAATTTAAACAACATATTTTATGATCGACAGAGTCAGC 192
Db 2318 ATACTACAGGTTTATTAATAATTTAGCTTTAGCTGAGTAATATATATTAATTAATGAATCG 2259

193 ATATGATATATGATTCAGAAATCGTTTTCGACGAGTTTCGAGTAGTAGTAGCCATTAATTT 252
Db 2258 ATATATTTCAAAAGGATCAGTTACTTTAGCTTCGATATATATAATTAATATACCATGATTT 2199

253 AATGTACATACATACTAATCGTAATAGTAGTATGATGAACATTTGTTATTTGTTATAAATA 312
Db 2198 TCAATTTTTTTTGTATATATCAAAAGTATTTCTTGMAAATTTGCAATTAATTTATCATTT 2139

313 TCCATAAAGACATCATGAAGACACATTTCTTTTCAGGGTCTGCAATTAATTTATGATACATTT 372
Db 2138 TCAATAATAATTTAAATTAATAACCAAAATTTTAAATAGTTTACTTAATGTTTAAATCTA 2079

373 CTAATAGAAACGAATTAATTAATTCAGTTGTAATGTAATGAATCTAATTTGA 421
Db 2078 TATTTACTAGTAAAGTTTAAACATGAATTAATTAATTAATTAATAGTAATAGA 2030

RESULT 12
US-09-790-988-1

Sequence 1, Application US/09790988
Patent No. 6632935

GENERAL INFORMATION:
APPLICANT: SHIGENOBU, SHUJI

APPLICANT: WATANABE, HIDEKI
APPLICANT: HAYTORI, MASAHIRA

APPLICANT: SAKAKI, YOSHIYUKI
TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS

FILE REFERENCE: 081356/0159
CURRENT APPLICATION NUMBER: US/09/790,988

CURRENT FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: JP2000-107160

PRIOR FILING DATE: 2000-04-07
NUMBER OF SEQ ID NOS: 7

SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1

LENGTH: 640681
TYPE: DNA

ORGANISM: Buchnera sp.
US-09-790-988-1

Query Match 3.2%; Score 46.6; DB 4; Length 640681;
Best Local Similarity 46.3%; Pred. No. 0.094;
Matches 183; Conservative 0; Mismatches 211; Indels 1; Gaps 1;

1 ATTATGATCTCAATACATACATATCTCATCTAGATCTAGATGTAATTTGTTATCATTTAAG 60
Db 383471 ATAAAGGTAAGAAAGTTATTTAAAGAAATATATTTAAAGAAATGTTACTTGAAGAAAG 383530

61 AAAGTTTTCAGCAATATGNNACGAAAAATGGCTACACTCGATGTAATTTGGTATCTCAAC 120
Db 383531 TACCATTATTAATAATATATAATTAATTAATGATCACACTGATTTAGGAGGTGTTTATA 383590

121 TCAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 180
Db 383591 AAAAATAAATAATTTATTTAATAATTTTATTAAGTAACGTAATTTATTTAAAAA 383650

181 GCAAGAGTCAGCATATGTAATTTGATTCAGATCGTTTTCAGCGAGTTCCGATGATAG 240
Db 383651 TTTATTTTATATATTTTATTAATAAATAAATAATATATTTAAATAAATAAATAAATAA 383710

QY 241 TAGCATATATTAATGACACTAATCGTGAATAGTATGATGATGAAACATTTGATCTT 300
DB 383711 TTTATATAAAAT-ATATAAATAATAATATTTATATAAAATATTTATATAAATTTTCAA 383769
QY 301 ATTGTATAAATATCCATAACACATCATGAAGACACTTTCTTCAGGGTCTGAATTAAT 360
DB 383770 ACTATTTTAATTTTATACATGTTTCATCTTTAACAATTTATATGATAAAAAACATTAAT 383829
QY 361 TATGATACAAATTTCTAATAGAAACGAATTAATTA 395
DB 383830 TATATAAATCAATTAAGAAAAAATTTAAAAATTA 383864

RESULT 13
US-08-526-964-2/c
; Sequence 2, Application US/08526964
; Patent No. 5698421
; GENERAL INFORMATION:
; APPLICANT: Lambowitz, Alan M
; APPLICANT: Zimmerly, Steven
; APPLICANT: Guo, Huatao
; APPLICANT: Yang, Jian
; TITLE OF INVENTION: Nucleotide Integrase Preparation
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calfee, Halter & Griswold
; STREET: 800 Superior Avenue
; CITY: Cleveland
; STATE: Ohio
; COUNTRY: USA
; ZIP: 44114
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/526,964
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Golrick, Mary E
; REGISTRATION NUMBER: 34,829
; REFERENCE/DOCKET NUMBER: 22727/00127
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2448 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
US-08-526-964-2

Query Match 3.2%; Score 46.2; DB 1; Length 2448;
Best Local Similarity 45.6%; Pred. No. 0.011;
Matches 159; Conservative 0; Mismatches 190; Indels 0; Gaps 0;
QY 73 AATAGNNACGACAAATGGCTACACTCGATGTAATGGTATCTCAACTCAACATTATAC 132
DB 2209 AATATACCTCTATGTAATGTTTAAACATGATGATTTCTACATCAATAGTTGAATTACAA 2150
QY 133 TTATACCAACATTAAGTTAGCAAAATTTAAACAACTATTTTATGATGATGCAAGATCAGC 192
DB 2149 ATACTACAGGTTTATTAATTAATAGCTTTAGCTGTAGGTAATATATTTTAAATGATCG 2090
QY 193 ATATGTAATTAATGATTCAGAAATCGTTTTCACGAGTTCGGATGATGATGATGATGATGAT 252
DB 2089 ATATATTTCAAAAGGATCAGTTACTTTTAGCTTCTGATATATATATAAATATACCATGATTT 2030

QY 253 AATGTACATACATATCGTGAATAGTATGATGATGAAACATTTGATCTTATGTAATAATA 312
DB 2029 TCAATTTTTTTTGAATATATCAAAAGTATTTCTTGGAAAAATTTGGCAATTAATTTATCATTT 1970
QY 313 TCCATAAACACATCATCAAGAACACACTTTCTTTCAGGGTCTGAAATTAATATGATACAAAT 372
DB 1969 TCAATATATTTTAATTTAATTAACCAATTTTATATAGTTTACTTATTTGTTTATCTA 1910
QY 373 CTATAGAAAAACGAATTAATTAAGTTGAAATTTGATGAAATCTTAATGA 421
DB 1909 TATTACTAGCTAAAGTTAATACACATGAATAATATAATACGTAAATAGA 1861

RESULT 14
US-08-946-617-2/c
; Sequence 2, Application US/08946617
; Patent No. 5869634
; GENERAL INFORMATION:
; APPLICANT: Lambowitz, Alan M
; APPLICANT: Zimmerly, Steven
; APPLICANT: Guo, Huatao
; APPLICANT: Yang, Jian
; TITLE OF INVENTION: Nucleotide Integrase Preparation
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calfee, Halter & Griswold
; STREET: 800 Superior Avenue
; CITY: Cleveland
; STATE: Ohio
; COUNTRY: USA
; ZIP: 44114
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/946,617
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Golrick, Mary E
; REGISTRATION NUMBER: 34,829
; REFERENCE/DOCKET NUMBER: 22727/00127
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2448 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
US-08-946-617-2

Query Match 3.2%; Score 46.2; DB 2; Length 2448;
Best Local Similarity 45.6%; Pred. No. 0.011;
Matches 159; Conservative 0; Mismatches 190; Indels 0; Gaps 0;
QY 73 AATAGNNACGACAAATGGCTACACTCGATGTAATGGTATCTCAACTCAACATTATAC 132
DB 2209 AATATACCTCTATGTAATGTTTAAACATGATGATTTCTACATCAATAGTTGAATTACAA 2150
QY 133 TTATACCAACATTAAGTTAGCAAAATTTAAACAACTATTTTATGATGATGCAAGATCAGC 192
DB 2149 ATACTACAGGTTTATTAATTAATAGCTTTAGCTGTAGGTAATATATTTTAAATGATCG 2090
QY 193 ATATGTAATTAATGATTCAGAAATCGTTTTCACGAGTTCGGATGATGATGATGATGATGAT 252
DB 2089 ATATATTTCAAAAGGATCAGTTACTTTTAGCTTCTGATATATATATAAATATACCATGATTT 2030

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2004, 03:16:30 ; Search time 1413 Seconds

(without alignments)
3762.369 Million cell updates/sec

Title: US-09-868-744B-1

Perfect score: 1444

Sequence: 1 atattatctcaatacatt.....gtactggaatgtagtatcc 1444

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2432557 seqs, 1840798884 residues

Total number of hits satisfying chosen parameters: 4865114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1184	82.0	1217	9	US-09-887-576-156
2	1174.8	81.4	1219	9	US-09-737-626A-9
3	1174.8	81.4	1219	14	US-10-427-169-9
4	1174.8	81.4	1219	14	US-10-427-180-9
5	1174	81.3	1742	9	US-09-737-626A-30
6	1174	81.3	1742	14	US-10-427-169-30
7	1174	81.3	1742	14	US-10-427-180-30
8	1104.6	76.5	1259	9	US-09-887-384A-5
9	1053	72.9	1202	9	US-09-887-384A-2
10	1001.8	69.4	1342	9	US-09-887-384A-6
11	950.2	65.8	1285	9	US-09-887-384A-3
12	784.4	54.3	910	9	US-09-887-384A-4
13	732.8	50.7	853	9	US-09-887-384A-1
14	144	10.0	1468	12	US-10-425-114-14693
15	94	6.5	94	9	US-09-887-384A-14

16	84.4	5.8	97	9	US-09-887-384A-12
17	79	5.5	79	9	US-09-887-384A-9
18	77.4	5.4	79	9	US-09-887-384A-17
19	76	5.3	2000	9	US-09-938-842A-3302
20	76	5.3	2000	11	US-09-938-842A-3302
21	75.4	5.2	1271	9	US-09-737-626A-10
22	75.4	5.2	1271	14	US-10-427-169-10
23	75.4	5.2	1271	14	US-10-427-180-10
24	73.8	5.1	1800	9	US-09-737-626A-29
25	73.8	5.1	1800	14	US-10-427-169-29
26	73.8	5.1	1800	14	US-10-427-180-29
27	65.4	4.5	67	9	US-09-887-384A-10
28	65.4	4.5	74	9	US-09-887-384A-7
29	63.4	4.4	77	9	US-09-887-384A-8
30	50.8	3.5	1612	9	US-09-770-445-3
31	50.6	3.5	3673778	14	US-10-312-841-1
32	50.4	3.5	6070	14	US-10-204-708-9
33	50.4	3.5	6070	14	US-10-311-455-213
34	50.2	3.5	71	9	US-09-887-384A-13
35	49.6	3.4	6255	14	US-10-311-455-934
36	49.2	3.4	2016	12	US-10-424-599-140494
37	49	3.4	158001	16	US-10-211-179-11
38	48.8	3.4	5798	14	US-10-311-455-1639
39	47.6	3.3	2444	12	US-10-424-599-19167
40	47.6	3.3	513509	10	US-09-754-853A-4
41	47.6	3.3	3673778	14	US-10-312-841-1
42	47.4	3.3	19087	14	US-10-311-455-766
43	47	3.3	63	9	US-09-887-384A-15
44	47	3.3	1760	12	US-10-424-599-80650
45	46.8	3.2	450	9	US-09-770-444-689

ALIGNMENTS

RESULT 1
US-09-887-576-156
; Sequence 156, Application US/09887576
; Patent No. US20020144047A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, P.
; APPLICANT: Brown, D.
; APPLICANT: Chang, H.
; APPLICANT: Zhu, T.
; APPLICANT: Han, B.
; APPLICANT: Wang, X.
; TITLE OF INVENTION: Promoters for regulation of plant expression
; FILE REFERENCE: 1360.001US1
; CURRENT APPLICATION NUMBER: US/09/887,576
; PRIOR FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: US 60/213,848
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/214,087
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/258,692
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 875
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 156
; LENGTH: 1217
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-887-576-156

Query Match 82.0%; Score 1184; DB 9; Length 1217;
Best Local Similarity 99.3%; Pred. No. 2.9e-272;
Matches 1211; Conservative 0; Mismatches 5; Indels 4; Gaps 2;
QY 164 CAACATATTTTATGATGCAAGTCAGATGATGATTAATTCAGATCGTTTTCAC 223
DB 1 CAACATATTTTATGATGCAAGTCAGATGATGATTAATTCAGATCGTTTTCAC 60

QY 224 GAGTTCGGATGTAGTAGCCATTATTTAATGTACATCTAATACTGTGAATGATGATG 283
Db 61 GAGTTCGGATGTAGTAGCCATTATTTAATGTACATCTAATACTGTGAATGATGATG 120
QY 284 ATGAAACATTGTATCTTATCTATAATATCCATAAACAACATCATGAAAGACATTTCTT 343
Db 121 ATGAAACATTGTATCTTATCTATAATATCCATAAACAACATCATGAAAGACATTTCTT 180
QY 344 TCAGGGTCTGAATTAATATGATACAAATTTAATAGAAAAGAAATTAATTAACGTGAAT 403
Db 181 TCAGGGTCTGAATTAATATGATACAAATTTAATAGAAAAGAAATTAATTAACGTGAAT 240
QY 404 TGATGAATCTAATTAATGAACCAACCAACGAGCACTAAGTTCCTGATGACT 463
Db 241 TGATGAATCTAATTAATGAACCAACCAACGAGCACTAAGTTCCTGATGACT 300
QY 464 CGGTTAAGTTAACTAAACCAACCAACGAGCACTAAGTTCCTGATGACT 523
Db 301 CGGTTAAGTTAACTAAACCAACCAACGAGCACTAAGTTCCTGATGACT 360
QY 524 ACAGTCATGAAGCACTAAGCAACCAACGAGCACTAAGTTCCTGATGACT 583
Db 361 ACAGTCATGAAGCACTAAGCAACCAACGAGCACTAAGTTCCTGATGACT 420
QY 584 TAAAAAATAGTTAAACCAACCAACGAGCACTAAGTTCCTGATGACT 642
Db 421 TAAAAAATAGTTAAACCAACCAACGAGCACTAAGTTCCTGATGACT 480
QY 643 TGTTGATGAAGTTTGTGTATAGATTCTTCGAAACCTTAGAATTTGTGTCGACGTTG 702
Db 1303 AACAGAAAGCTATTCTCTGATTCATCAGGGGTTTATTTGACTGTATTGAACCTTTTGTG 1362

QY 1138 AACAGAAAGCTATTCTCTGATTCATCAGGGGTTTATTTGACTGTATTGAACCTTTTGTG 1197
QY 1363 TGTTGAGCTCATAAAAA 1382
Db 1198 TGTTGAGCTCATAAAAA 1217

RESULT 2
US-09-737-626A-9
; Sequence 9. Application US/09737626A
; Patent No. US20020144304A1
; GENERAL INFORMATION:
; APPLICANT: Fincher, Karen
; APPLICANT: Flaisinski, Stanislaw
; APPLICANT: Wilkinson, Jack
; TITLE OF INVENTION: No. US20020144304A1el Plant Expression Constructs
; FILE REFERENCE: 38-21(51499)B
; CURRENT APPLICATION NUMBER: US/09737, 626A
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 09/737, 626
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 1219
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1) .. (1219)
; OTHER INFORMATION: Act2 promoter: polynucleotide sequence and intron
US-09-737-626A-9

Query Match 81.4%; Score 1174.8; DB 9; Length 1219;
Best Local Similarity 98.7%; Pred. No. 4.6e-270;
Matches 1206; Conservative 0; Mismatches 12; Indels 4; Gaps 2;
QY 164 CAACATTTTTATGTATGCAAGAGTCAGCATATGTATTAATGATTCAGAAATCGTTTTCAC 223
Db 1 CAACATTTTTATGTATGCAAGAGTCAGCATATGTATTAATGATTCAGAAATCGTTTTCAC 60
QY 224 GAGTTCGGATGTAGTAGCCATTATTTAATGTACATCTAATACTGTGAATGATGATG 283
Db 61 GAGTTCGGATGTAGTAGCCATTATTTAATGTACATCTAATACTGTGAATGATGATG 120
QY 284 ATGAAACATTGTATCTTATCTATAATATCCATAAACAACATCATGAAAGACATTTCTT 343
Db 121 ATGAAACATTGTATCTTATCTATAATATCCATAAACAACATCATGAAAGACATTTCTT 180
QY 344 TCAGGGTCTGAATTAATATGATACAAATTTAATAGAAAAGAAATTAATTAACGTGAAT 403
Db 181 TCAGGGTCTGAATTAATATGATACAAATTTAATAGAAAAGAAATTAATTAACGTGAAT 240
QY 404 TGATGAATCTAATTAATGAACCAACCAACGAGCACTAAGTTCCTGATGACT 463
Db 241 TGATGAATCTAATTAATGAACCAACCAACGAGCACTAAGTTCCTGATGACT 300
QY 464 CGGTTAAGTTAACTAACAATAAAAAACGAGCTGTATGACACGCGGATCGAGCAGGTC 523
Db 301 CGGTTAAGTTAACTAACAATAAAAAACGAGCTGTATGACACGCGGATCGAGCAGGTC 360
QY 524 ACAGTCATGAAGCACTAAGCAACCAACGAGCACTAAGTTCCTGATGACT 583
Db 361 ACAGTCATGAAGCACTAAGCAACCAACGAGCACTAAGTTCCTGATGACT 420
QY 584 TAAAAAATAGTTAAACCAACCAACGAGCACTAAGTTCCTGATGACT 642
Db 421 TAAAAAATAGTTAAACCAACCAACGAGCACTAAGTTCCTGATGACT 480
QY 643 TGTTGATGAAGTTTGTGTATAGATTCTTCGAAACCTTAGAATTTGTGTCGACGTTG 702
Db 481 TGTTGATGAAGTTTGTGTATAGATTCTTCGAAACCTTAGAATTTGTGTCGACGTTG 540

Db	1253	TGAACTCTCCCTCAACTCA---TCTCTCCCGCTCTTCTCTTCCAAGGTAATAGGAACCTT	1309
QY	943	CTCGAATCTACTTTAATTTGCTGGATCTCGAATCTTGTTTCTCAATTTCCCTTGAGATCTGG	1002
Db	1310	CTCGAATCTACTTTAATTTGCTGGATCTCGAATCTTGTTTCTCAATTTCCCTTGAGATCTGG	1369
QY	1003	AATTCGTTTAAATTTGGATCTGTGAACCTCCACCTAAATCTTTTGGTTTTTACTAGAAATCGAT	1062
Db	1370	AATTCGTTTAAATTTGGATCTGTGAACCTCCACCTAAATCTTTTGGTTTTTACTAGAAATCGAT	1429
QY	1063	CTAAGTTGACCGATCAGTTAGCTCGATTAGCTACCAAGAAATTTGGCTTTGACCTTGATGG	1122
Db	1430	CTAAGTTGACCGATCAGTTAGCTCGATTAGCTACCAAGAAATTTGGCTTTGACCTTGATGG	1489
QY	1123	AGAGATCCATGTTTCATGTTTACCTCGGAAATGATTTGTATATGTGAATTCGAATCTGAACT	1182
Db	1490	AGAGATCCATGTTTCATGTTTACCTCGGAAATGATTTGTATATGTGAATTCGAATCTGAACT	1549
QY	1183	GTTGAAGTTAGATTTGAATCTGAACACCTGCTCAATGTTAGATTTGAATCTGAACACCTGTTTAA	1242
Db	1550	GTTGAAGTTAGATTTGAATCTGAACACCTGCTCAATGTTAGATTTGAATCTGAACACCTGTTTAA	1609
QY	1243	GTTAGATGAGTTTGTGTAAGATCTTCGAAACCTTAGAATTTGTAGTGTGCTGACGTTG	1302
Db	1610	GTTAGATGAGTTTGTGTAAGATCTTCGAAACCTTAGAATTTGTAGTGTGCTGACGTTG	1669
QY	1303	AACAGAAAGCTATTTCTGATTCATCAAGGCTTTTATTTGACCTGATTTGAATCTTTTGTG	1362
Db	1670	AACAGAAAGCTATTTCTGATTCATCAAGGCTTTTATTTGACCTGATTTGAATCTTTTGTG	1729
QY	1363	TGTTTGCAGC	1372
Db	1730	TGTTTGCAGC	1739

RESULT 6

US-10-427-169-30

; Sequence 30, Application US/10427169

; Publication No. US20030199681A1

; GENERAL INFORMATION:

; APPLICANT: Fincher, Karen

; APPLICANT: Flaszinski, Stanislaw

; APPLICANT: Wilkinson, Jack

; TITLE OF INVENTION: No. US20030199681A1 Plant Expression Constructs

; FILE REFERENCE: 38-21(51499)B

; CURRENT APPLICATION NUMBER: US/10/427,169

; CURRENT FILING DATE: 2003-05-01

; PRIOR APPLICATION NUMBER: US/09/737,626A

; PRIOR FILING DATE: 2002-02-25

; PRIOR APPLICATION NUMBER: 09/737,626

; PRIOR FILING DATE: 2000-12-15

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 30

; LENGTH: 1742

; TYPE: DNA

; ORGANISM: artificial sequence

; FEATURE:

; NAME/KEY: promoter

; LOCATION: (1)..(1742)

; OTHER INFORMATION: chimeric promoter fusion CamV and Act2 polynucleotides +

; OTHER INFORMATION: tto

US-10-427-169-30

Query Match 81.3%; Score 1174; DB 14; Length 1742;

Best local Similarity 99.3%; Pred. No. 8.5e-270;

Matches 1201; Conservative 0; Mismatches 5; Indels 4; Gaps 2;

QY	164	CAACTATTTTATGATGATCAAGAGTCAGCATATGTAATAATTTGATTCAGAAATCGTTTTGAC	223
Db	533	CAACTATTTTATGATGATCAAGAGTCAGCATATGTAATAATTTGATTCAGAAATCGTTTTGAC	592
QY	224	GAGTTCGGATGTAGTAGTAGCCATATTTAATGTACATATACTAATCTGTAATGATGATG	283

Db 593 GAGTTCGGATCTAGTAGGAGCCATTAATTAATGTACATACATAATCGTAATGAGTAG 652
Qy |||||
Db 284 ATGAACATGTTATCTTATGTAATAATATCCATAAACACATCATGAAGACATTTCTT 343
Qy |||||
Db 653 ATGAACATGTTATCTTATGTAATAATATCCATAAACACATCATGAAGACATTTCTT 712
Qy |||||
Db 344 TCAGGCTCTGAATTAATATGATACAAATCTTAATAGAAACGAATTAATTAATGTAAT 403
Qy |||||
Db 713 TCAGGCTCTGAATTAATATGATACAAATCTTAATAGAAACGAATTAATTAATGTAAT 772
Qy |||||
Db 404 TGTATGAATCTAAATTTGAACAAGCAACCAACGAGGAGCTAAACGTTGCTGGAATGACT 463
Qy |||||
Db 773 TGTATGAATCTAAATTTGAACAAGCAACCAACGAGGAGCTAAACGTTGCTGGAATGACT 832
Qy |||||
Db 464 CGGTTTAAGTTAAACCACTAAACCAACGAGGAGCTGATGTAACACGCGGATCGACGGTC 523
Qy |||||
Db 833 CGGTTTAAGTTAAACCACTAAACCAACGAGGAGCTGATGTAACACGCGGATCGACGGTC 892
Qy |||||
Db 524 ACAGTCATGAGGCTAAACCAACGAGGAGCTGATGTAACACGCGGATCGACGGTC 583
Qy |||||
Db 893 ACAGTCATGAGGCTAAACCAACGAGGAGCTGATGTAACACGCGGATCGACGGTC 952
Qy |||||
Db 584 TAAATAATGTTAAACCACTAAACCAACGAGGAGCTGATGTAACACGCGGATCGACGGTC 642
Qy |||||
Db 953 TAAATAATGTTAAACCACTAAACCAACGAGGAGCTGATGTAACACGCGGATCGACGGTC 1012
Qy |||||
Db 643 TGTGGTCTGAATGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 702
Qy |||||
Db 1013 TGTGGTCTGAATGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1072
Qy |||||
Db 703 AGTTGTAAGAGTAACCCGCTATATAATTAATTAATTAATTAATTAATTAATTAATTAAT 762
Qy |||||
Db 1073 AGTTGTAAGAGTAACCCGCTATATAATTAATTAATTAATTAATTAATTAATTAATTAAT 1132
Qy |||||
Db 763 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 822
Qy |||||
Db 1133 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1192
Qy |||||
Db 823 AACAGGAAGAGACTAAGAGAGAAAGTAAGAGTAATCCAGGAGATTCATCTCCGTTT 882
Qy |||||
Db 1193 AACAGGAAGAGACTAAGAGAGAAAGTAAGAGTAATCCAGGAGATTCATCTCCGTTT 1252
Qy |||||
Db 883 TGAATCTCTCAATCTCATCT 942
Qy |||||
Db 1253 TGAATCTCTCAATCTCAATCT 1309
Qy |||||
Db 943 TCTGATCTACTTTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1002
Qy |||||
Db 1310 TCTGATCTACTTTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1369
Qy |||||
Db 1003 AATTGCTTTAATTTGATCTGTAACCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 1062
Qy |||||
Db 1370 AATTGCTTTAATTTGATCTGTAACCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 1429
Qy |||||
Db 1063 CTAAGTTGACCGATCAGTTAGCTGATTTAGCTACAGAAATTTGCTTGAATCTGATG 1122
Qy |||||
Db 1430 CTAAGTTGACCGATCAGTTAGCTGATTTAGCTACAGAAATTTGCTTGAATCTGATG 1489
Qy |||||
Db 1123 AGAGATCCATGTTTATGTTTACCTGGGAATGATTTGATATGTAATGTAATCTGAATCTGAAT 1182
Qy |||||
Db 1490 AGAGATCCATGTTTATGTTTACCTGGGAATGATTTGATATGTAATGTAATCTGAATCTGAAT 1549
Qy |||||
Db 1183 GTTAGGTTAGATTGAATCTGAACATGCTGATGATGATGATGATGATGATGATGATGATGAT 1242
Qy |||||
Db 1550 GTTAGGTTAGATTGAATCTGAACATGCTGATGATGATGATGATGATGATGATGATGATGAT 1609
Qy |||||
Db 1243 GTTAGGTTAGATTGAATCTGAACATGCTGATGATGATGATGATGATGATGATGATGATGAT 1302
Qy |||||
Db 1610 GTTAGGTTAGATTGAATCTGAACATGCTGATGATGATGATGATGATGATGATGATGATGAT 1669
Qy |||||
Db 1303 AACAGGAAGCTATTTCTGATTAATCAGGCTTATTTGATGATTAATGATCTTTTGTG 1362
Qy |||||

Db 1670 AACAGGAAGCTATTTCTGATTCATCAGGGTTATTTGCTGATGATGAACCTCTTTTGTG 1729
Qy |||||
Db 1363 TGTGTCAGC 1372
Qy |||||
Db 1730 TGTGTCAGC 1739
Qy |||||

RESULT 7
US-10-427-180-30
; Sequence 30, Application US/10427180
; Publication No. US20030199682A1
; GENERAL INFORMATION:
; APPLICANT: Fincher, Karen
; APPLICANT: Flasiński, Stanisław
; APPLICANT: Wilkinson, Jack
; TITLE OF INVENTION: No. US20030199682A1 Plant Expression Constructs
; FILE REFERENCE: 38-21(51499)B
; CURRENT APPLICATION NUMBER: US/10/427,180
; CURRENT FILING DATE: 2003-05-01
; PRIOR APPLICATION NUMBER: US/09/737,626A
; PRIOR FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 09/737,626
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 30
; LENGTH: 1742
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1742)
; OTHER INFORMATION: chimeric promoter fusion CaMv and Act2 polynucleotides + Act2
; OTHER INFORMATION: tfo
US-10-427-180-30

Query Match 81.3%; Score 1174; DB 14; Length 1742;
Best Local Similarity 99.3%; Pred. No. 8.5e-270;
Matches 1201; Conservative 0; Mismatches 5; Indels 4; Gaps 2;

Qy 164 CAACATATTTTATGTAATGCAAGAGTCAGCATATGTAATTAATGATTCAGAAATCGTTTGTAC 223
Db 533 CAACATATTTTATGTAATGCAAGAGTCAGCATATGTAATTAATGATTCAGAAATCGTTTGTAC 592
Qy 224 GAGTTCGGATGTAAGTAGAGCCATTAATTAATGTAACATCTAATCGTGAATAGTATG 283
Db 593 GAGTTCGGATGTAAGTAGAGCCATTAATTAATGTAACATCTAATCGTGAATAGTATG 652
Qy 284 ATGAAACATTTGATCTTATTTGTAATAATATCCATATAACACATCATGAAAGACATTTCTT 343
Db 653 ATGAAACATTTGATCTTATTTGTAATAATATCCATATAACACATCATGAAAGACATTTCTT 712
Qy 344 TCAGGCTCTGAATTAATTAATGATACAAATCTTAATAGAAACGAATTAATTAATGATG 403
Db 713 TCAGGCTCTGAATTAATTAATGATACAAATCTTAATAGAAACGAATTAATTAATGATG 772
Qy 404 TGTATGAATCTAAATTTGAACAAGCAACCAACGAGGAGCTAAACGTTGCTGGAATGACT 463
Db 773 TGTATGAATCTAAATTTGAACAAGCAACCAACGAGGAGCTAAACGTTGCTGGAATGACT 832
Qy 464 CGGTTTAAGTTAAACCACTAAACCAACGAGGAGCTGATGTAACACGCGGATCGACGGTC 523
Db 833 CGGTTTAAGTTAAACCACTAAACCAACGAGGAGCTGATGTAACACGCGGATCGACGGTC 892
Qy 524 ACAGTCATGAGGCTAAACCAACGAGGAGCTGATGTAACACGCGGATCGACGGTC 583
Db 893 ACAGTCATGAGGCTAAACCAACGAGGAGCTGATGTAACACGCGGATCGACGGTC 952
Qy 584 TAAATAATGTTAAACCACTAAACCAACGAGGAGCTGATGTAACACGCGGATCGACGGTC 642
Db 953 TAAATAATGTTAAACCACTAAACCAACGAGGAGCTGATGTAACACGCGGATCGACGGTC 1012
Qy 643 TGTGCTGAAATGATTCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 702

Db 1004 GTTGAAGTAGATATTGACCTGTAACCTGCTCAATGTAGATATTGACCTGTAACCTGTTAA 1063
Qy 1243 GTTAGAGAAAGTTGTGTATAGATCTTCTGAAACCTTAGAGATTTGAGTGTGACGTTG 1302
Db 1064 GTTAGATGAAGTTTGTGTATAGATCTTCTGAAACCTTAGAGATTTGAGTGTGACGTTG 1123
Qy 1303 AACAGAAGCTATTCTGATCAATCAGGCTTTTATTTGACTGTATTTGAACTCTTTTGTG 1362
Db 1124 AACAGAAGCTATTCTGATCAATCAGGCTTTTATTTGACTGTATTTGAACTCTTTTGTG 1183
Qy 1363 TGTGTGAGCTCATAAAAATGCTGAGGCTGACGATATTCAACCAATCGTGTGACAA 1422
Db 1184 TGTGTGAGCTCATAAAAATGCTGAGGCTGACGATATTCAACCAATCGTGTGACAA 1243
Qy 1423 TGTGTGAGCTCATAAAAATGCTGAGGCTGACGATATTCAACCAATCGTGTGACAA 1437
Db 1244 TGTGTGAGCTCATAAAAATGCTGAGGCTGACGATATTCAACCAATCGTGTGACAA 1258

RESULT 9
US-09-887-384A-2
; Sequence 2, Application US/09887384A
; Patent No. US2002004992A1
; GENERAL INFORMATION:
; APPLICANT: Hamilton, Carol
; TITLE OF INVENTION: MODIFIED PROMOTERS
; FILE REFERENCE: actin promoters
; CURRENT APPLICATION NUMBER: US/09/887,384A
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1202
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Actin based promoter
US-09-887-384A-2

Query Match 72.9%; Score 1053; DB 9; Length 1202;
Best Local Similarity 94.0%; Pred. No. 5.4e-241;
Matches 1144; Conservative 0; Mismatches 55; Indels 18; Gaps 4;
Qy 164 CAACATTTTATGATGCAAGAGTCAGCATATGTATTAATTTGATTCAGAACTGTTTGTAC 223
Db 1 CAACATTTTATGATGCAAGAGTTTCATCTTGTATTAATTTGATTCAGAACTGTTTGTAC 60
Qy 224 GAGTTCGAGTAGTAGTAGGCAATTTAATGTATACATTAATTCGTGAATAGTAGATG 283
Db 61 GAGTTCGAGTAGTAGTAGGCAATTTAATGTATACATTAATTCGTGAATAGTAGATG 120
Qy 284 ATGAACATGTATCTTATGTATATATCCATTAACATCATGCAAGACATCTTCT 343
Db 121 ATGAACATGTATCTTATGTATATATCCATTAACATCATGCAAGACATCTTCT 180
Qy 344 TCAGGCTCTGAATTAATATGATACATTAATTAAGAAACGAATTAATTAACGTTGAAT 403
Db 181 TCAGGCTCTGAATTAATATGATACATTAATTAAGAAACGAATTAATTAACGTTGAAT 240
Qy 404 TGTATGAATCTTAATTAAGAAACGAATTAATTAAGAAACGAATTAATTAACGTTGAAT 463
Db 241 TGTATGAATCTTAATTAAGAAACGAATTAATTAAGAAACGAATTAATTAACGTTGAAT 300
Qy 464 CGGTTAAGTAAACCACTAAACCAACGAGCTGTATGTAACGAGGCTGACGAGGTC 523
Db 301 AAGTTAAGTAAACCACTAAACCAACGAGCTGTATGTAACGAGGCTGACGAGGTC 360
Qy 524 ACAGTCAATGAAGCCATCAACCAACGAACCTAATCAAGGCTGAGATGATTAATGATT 583
Db 361 ACAGTCAATGAAGCCATCAACCAACGAACCTAATCAAGGCTGAGATGATTAATGATT 420

Qy 584 TAAAAATTAGTTAAACGAGGAAAA-GCTGTCTGACAGCGAGTCACTTATCTTTACC 642
Db 421 TAAAAATTAGTTAAACGAGGAAAAAGCTGTCTGACAGCGAGTCACTTATCTTTACC 480
Qy 643 TGTGCTCGAAATGATTCGTGTCTGTGATTTTAAATTTTAAATTTTAAAGGCGGAAAAATA 702
Db 481 TGTGCTCGAAATGATTCGTGTCTGTGATTTTAAATTTTAAATTTTAAAGGCGGAAAAATA 540
Qy 703 AGTTGTAAGAGATAAACCGGCTATATAATTCATATATTTCTCCCGCTTTGAATTTG 762
Db 541 AGTTGTAAGAGATAAACCGGCTATATAATTCATATATTTCTCCCGCTTTGAATTTG 600
Qy 763 TCTGTTGTCT 822
Db 601 TCTGTTGTCT 660
Qy 823 AACAGAAGAGAGCTAAG 882
Db 661 AACAGAAG 706
Qy 883 TGAATCT 942
Db 707 TGAATCT 763
Qy 943 TCTGATCTACTTTTATTTTGTCTGATCTCGATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1002
Db 764 TCTGATCTACTTTTATTTTGTCTGATCTCGATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 823
Qy 1003 AATTCGTTTAAATTTGAGTGTGAACTCCACTAAATCTTTTGTCTTGTCTTGTCTTGTCT 1062
Db 824 TTTTCGTTTAAATTTGAGTGTGAAATTTAGATCCTAAATCTTTTGTCTTGTCTTGTCTTGT 883
Qy 1063 CTAGTTGACCGGCTAGTTAGTCTGATTTAGTCTACAGAAATTTGCTTGTCTTGTCTTGTCT 1122
Db 884 CTAGTTGACCGGCTAGTTAGTCTGATTTAGTCTACAGAAATTTGCTTGTCTTGTCTTGTCT 943
Qy 1123 AGAGATCCATGTTTCATGTTTACCTCGGAAATGATTTGTATATGTGAATTTGAATCTGAAT 1182
Db 944 AGAGATCCATGTTTCATGTTTACCTCGGAAATGATTTGTATATGTGAATTTGAATCTGAAT 1003
Qy 1183 GTTGAAGTTAGTTGATTTGATAGATTTCTCGAAACCTTAGGATTTGTAGTGTCTGACGTTG 1242
Db 1004 GTTGAAGTTAGTTGATTTGATAGATTTCTCGAAACCTTAGGATTTGTAGTGTCTGACGTTG 1063
Qy 1243 GTTGAAGTTAGTTGATTTGATAGATTTCTCGAAACCTTAGGATTTGTAGTGTCTGACGTTG 1302
Db 1064 GTTGAAGTTAGTTGATTTGATAGATTTCTCGAAACCTTAGGATTTGTAGTGTCTGACGTTG 1123
Qy 1303 AACAGAAGCTATTCTGATCAATCAGGCTTTTATTTGACTGTATTTGAACTCTTTTGTG 1362
Db 1124 AACAGAAGCTATTCTGATCAATCAGGCTTTTATTTGACTGTATTTGAACTCTTTTGTG 1183
Qy 1363 TGTGTGAGCTCATAAAA 1379
Db 1184 TGTGTGAGCTCATAAAA 1200

RESULT 10
US-09-887-384A-6
; Sequence 6, Application US/09887384A
; Patent No. US2002004992A1
; GENERAL INFORMATION:
; APPLICANT: Hamilton, Carol
; TITLE OF INVENTION: MODIFIED PROMOTERS
; FILE REFERENCE: actin promoters
; CURRENT APPLICATION NUMBER: US/09/887,384A
; CURRENT FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1342
; TYPE: DNA
; ORGANISM: Artificial Sequence

451 GCCTGATTCAGTCTCGTTTAACTTAACCACTAAATAAAGCGAGCTGTCACTACACGG 510
Db |
312 GCCTGGAATGAAGAAAGTTTAAAGTAACCACTAAATAAAGCGAGCTGTCACTACACGG 371
Qy |
511 GATCGAGAGGTTCACAGT-----CATGAAGCCATCAAGCAAAAGAACTAATCAAGGGG 565
Db |
372 GATCGAGAGGTTCACAGTCAATCCATGAAGCCATCAAGCAAAAGAACTAATCAAGGGG 431
Qy |
566 TGAGATGATTAATTAAGTTTAAATTAAGTTTAAACAGAGGGAATAA-GCTGTCTGACAGCCA 624
Db |
432 TGAGATGATTAATTAAGTTTAAATTAAGTTTAAACAGAGGGAATAA-GCTGTCTGACAGCCA 491
Qy |
625 GGTCAAGTTA---TCTTTACCTGTGTGGAATGATTCGTCTGTCGATTTTAAATTAAT 681
Db |
492 GGTCAAGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 551
Qy |
682 T-----TTTTGAAGCCGCAAAATAAAGTTTGAAGATTAAGAAACCCGCTCATATAA 731
Db |
552 TCCAAATTAATTAATTAAGGCGCAAAATAAAGTTTGAAGATTAAGAAACCCGCTCATATAA 611
Qy |
732 ATTCAAT-----ATATTTCTCTCCGCTTTTGAATTTCTCTGTTCTCTCTCACTT 782
Db |
612 ATTCATCAACCTCGATATTTCTCTCCGCTTTTGAATTTCTCTGTTCTCTCTCACTT 671
Qy |
783 TCATCAGCGCTTTTGAATCTCCGCGACTTGACAGAGAAACAGAGGAAGAACTAAGA 842
Db |
672 TCATCAGCGCTTTTGAATCTCCGCGACTTGACAGAGAAACAGAGGAAGAACTAAGA 731
Qy |
843 GAGAAATGAAGATTAATCAAGAGATTCATTCGCTGTTTGAATCTTCCTCAATCTCAT 902
Db |
732 GAGAAATGAAGATTAATCAAGAGATTCATTCGCTGTTTGAATCTTCCTCAATCTCAT- 790
Qy |
903 CTTCCTCTCCGCTTTTCTTCCAGGTAATAGCACTTTC-TGGATCTACTTTATTTG 961
Db |
791 --TCTTCTCCGCTTTTCTTCCAGGTAATAGCACTTTC-TGGATCTACTTTATTTG 848
Qy |
962 CTGATCTCGATCTTGTCTTCTCAATTTCTTGAATCTGGAATTCG-TTAAATTTGA 1019
Db |
849 CTGATCTCGATCTTGTCTTCTCAATTTCTTGAATCTGGAATTCG-TTAAATTTGA 908
Qy |
1020 TCTGTGAACCTCGACTAAATCTTTTGTGTTTACTAGATCGATCTAAGTTCACCGATCAG 1079
Db |
909 TCGAATTTAGATCACTAAATCTTTTGTGTTTACTAGATCGATCTAAGTTCACCGATCAG 968
Qy |
1080 TTAGCTGATTAATAGTACAGAAATTTGGCTTGACCTTTGATGAGA---GATCCATCTTC 1136
Db |
969 TTAGCTGATTAATAGTACAGAAATTTGGCTTGACCTTTGATGAGA---GATCCATCTTC 1028
Qy |
1137 ATGTTACCTGGGAAATGATTTGTATATGTGAATTTGAATCTGA---ACTGTTGAAGTTAGA 1194
Db |
1029 ATGTTACCTGGGAAATGATTTGTATATGTGAATTTGAATCTGA---ACTGTTGAAGTTAGA 1088
Qy |
1195 TTGAATCTGAACACTGTCAATGTTTGAATTTGAATCTGAACACTGTTTAAGTTAGATGAAGT 1254
Db |
1089 TTATGATCTGAACACTGTCAATGTTTGAATTTGAATCTGAACACTGTTTAAGTTAGATGAAGT 1148
Qy |
1255 TTGTGTATAGATTTCTGAAACCTTAGGATTTGTAGTCTGATGTTGAAACAGAAAGCTA 1314
Db |
1149 TTGTGTATAGATTTCTGAAACCTTAGGATTTGTAGTCTGATGTTGAAACAGAAAGCTA 1208
Qy |
1315 TTTCGTATCAATCAGGGTT-----TATTTGATGATTAAGTCAATCTTTTGTGTTG 1364
Db |
1209 TTTCGTATCAATCAGGGTTTTCGGTTTCTATTTGATGATTAAGTCAATCTTTTGTGTTG 1268
Qy |
1365 TTTCGAGCTCATAAA 1379
Db |
1269 TTTCGAGCTCATAAA 1283

RESULT 12
US-09-867-384A-4
; Sequence 4, Application US/09887384A

; Patent No. US2002004992A1
; GENERAL INFORMATION:
; APPLICANT: Hamilton, Carol
; TITLE OF INVENTION: MODIFIED PROMOTERS
; FILE REFERENCE: actin promoters
; CURRENT APPLICATION NUMBER: US/09/887,384A
; CURRENT FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO. 4
; LENGTH: 910
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Actin based
; OTHER INFORMATION: promoter
US-09-887-384A-4

Query Match 54.3%; Score 784.4; DB 9; Length 910;
Best Local Similarity 93.6%; Pred. No. 6.1e-177;
Matches 867; Conservative 0; Mismatches 41; Indels 18; Gaps 4;
Qy 513 TCGAGCAGGTTCACAGTCAATGAAGCCATCAAGCAAAAGAACTAATCCAGGGGTGAGATG 572
Db 1 TCGAGCAGGTTCACAGTCAATGAAGCCATCAAGCAAAAGAACTAATCCAGGGGTGAGATG 60
Qy 573 ATTAATTAAGTTTAAATAATTAAGTTTAAACAGAGGGAATAA-GCTGTCTGACAGCCAGTCAAG 631
Db 61 ATTAATTAAGTTTAAATAATTAAGTTTAAACAGAGGGAATAA-GCTGTCTGACAGCCAGTCAAG 120
Qy 632 TTATCTTTACCTGTGTGTGCGAATGATTCGTGTCTGATTTTAAATTTTAAATTTTGAAG 691
Db 121 TTATCTTTACCTGTGTGTGCGAATGATTCGTGTCTGATTTTAAATTTTAAATTTTGAAG 180
Qy 692 GCCGAAATAAAGTTTGAAGATTAACCCGCTATATAAATTCATATATTTCTCTCCC 751
Db 181 GCCGAAATAAAGTTTGAAGATTAACCCGCTATATAAATTCATATATTTCTCTCCC 240
Qy 752 GCTTTGAATTTGTCTGTTCT 811
Db 241 GCTTTGAATTTGTCTGTTCT 300
Qy 812 TCACAGAGAAACAGAGAAAGAACTAAGAGAAAGTAAGAGATAATCCAGGAGATTC 871
Db 301 TCACAGAGAAACAGAGAAAG---AGAGAGAAAGTAAGAGATAATCCAGG--- 347
Qy 872 ATTCCTCGTTTGAATCT 931
Db 348 -TTCTCTCGTTTGAATCT 403
Qy 932 AATAGGAACCTTCTGGAATCTACTTTATTTGCTGGAATCTGATCTGTTTCTCAATTTCC 991
Db 404 AATAGGAACCTTCTGGAATCTACTTTATTTGCTGGAATCTGATCTGTTTCTCAATTTCC 463
Qy 992 TTGAGATCTGGAATCTGTTTAAATTTGAATCTGGAACCTCCACTTAATCTTTTGGTTTAA 1051
Db 464 TTGAGATCTGTTTAAATTTGAATTTGAATTTGAATTTAGATCTTAATCTTTTGGTTTAA 523
Qy 1052 CTAGAATCTGATCTAAGTTGACCGGATCAGTTAGTCTGATTTAGTACCAGAAATTTGGCTT 1111
Db 524 CTAGAATCTGATCTAAGTTGACCGGATCAGTTAGTCTGATTTAGTACCAGAAATTTGGCTT 583
Qy 1112 GACCTTGATCGAGAGATCCATGTTTCATGTTTACCTGGAAATGATTTGTATATGTAATG 1171
Db 584 GACCTTGATCGAGAGATCCATGTTTCATGTTTACCTGGAAATGATTTGTATATGTAATG 643
Qy 1172 AAATCTGAATCTGTTGAGTTAGATTTGAATCTGAACTCACTGTCATGTTAGATTTAGT 1231
Db 644 AAATCTGAATCTGTTGAGTTAGATTTGAATCTGAACTCACTGTCATGTTAGATTTAGT 703
Qy 1232 ACACCTGTTTAAAGTTAGATGAAGTTTGTATAGATTTCTTCCAAACCTTAGGATTTGATG 1291
Db 704 TACTGTTTAAAGTTAGATGAAGTTTGTATAGATTTCTTCCGAACTTAGGATTTGATG 763

QY 1292 GTCGATCGTTGACAGAAAGCTATTCTGATCAATCAGGGTTATTGACTGTATTGAA 1351
DB 764 GTCGATCGTTGACAGAAAGCTATTCTGATCAATCAGGGTTATTGACTGTATTGAA 823
QY 1352 CTCCTTTTGTGTTGTCAGCTCATAAAAATGGCTGAGGCTCAGCATATTCAACCAATC 1411
DB 824 CTCCTTTTGTGTTGTCAGCTCATAAAAATGGCTGAGGCTCAGCATATTCAACCAATC 983
QY 1412 GTGTGTGACAAATGGTACTGGAATGGT 1437
DB 884 GTGTGTGACAAATGGTACTGGAATGGT 909

RESULT 13

US-09-887-384A-1

; Sequence 1, Application US/09887384A

; Patent No. US2002004992A1

; GENERAL INFORMATION:

; APPLICANT: Hamilton, Carol

; TITLE OF INVENTION: MODIFIED PROMOTERS

; FILE REFERENCE: actin promoters

; CURRENT APPLICATION NUMBER: US/09/887,384A

; CURRENT FILING DATE: 2001-06-22

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 853

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: promoter

US-09-887-384A-1

Query Match 50.7%; Score 732.8; DB 9; Length 853;
Best Local Similarity 93.7%; Pred. No. 1.2e-164;
Matches 813; Conservative 0; Mismatches 37; Indels 18; Gaps 4;

QY 513 TCGAGCAGGTCACAGTCATGAGCCATCAAGCAAGAACTAATCCAGGGGTGAGATG 572
DB 1 TCGAGCAGGTCACAGTCATGAGCCATCAAGCAAGAACTAATCCAGGGGTGAGATG 60
QY 573 ATTAATAGTTTAAAAATAGTTAAACAGAGGAAAA-GCTGTCTGACAGCCAGTCACG 631
DB 61 ATTAATAGTTTAAAAATAGTTAAACAGAGGAAAAAGCTGTCTGACAGCCAGTCACG 120
QY 632 TTATCTTTACCTGTGTGCGAAATGANTGCTGTCTGCAATTTAATTTATTTTGAAG 691
DB 121 TTATCTTTACCTGTGTGCGAAATGANTGCTGTCTGCAATTTAATTTATTTTGAAG 180
QY 692 GCCGAAATAAGTTGTAAGAGATAAACCGCTATATAATTCATATATTTCTCC 751
DB 181 GCCGAAATAAGTTGTAAGAGATAAACCGCTATATAATTCATATATTTCTCTCC 240
QY 752 GCTTTGAATTTGCTGCT 811
DB 241 GCTTTGAATTTGCTGCT 300
QY 812 TGACAGAGAGAACAGATC 871
DB 301 TGACAGAGAGAGAACAGATC 347
QY 872 ATTCTCGGTTTGAATTTCTGCAATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 931
DB 348 -TTCTCGGTTTGAATTTCTGCAATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 403
QY 932 AATAGAACTTTCTGGAATCTACTTTATTTTCTGGAATCTGGAATCTGGAATCTGGAATCT 991
DB 404 AATAGAACTTTCTGGAATCTACTTTATTTTCTGGAATCTGGAATCTGGAATCTGGAATCT 463
QY 992 TTGAGATCTGGAATCTGTTTAAATTTGGAATCTGTAACCTCCATAATCTTTTGGTTTAA 1051
DB 464 TTGAGATCTGTTTAAATTTGGAATCTGTAACCTCCATAATCTTTTGGTTTAA 523

QY 1052 CTAGATCGATCTAAGTTGACCGATCAGTTAGCTGATATAGCTACCAAGATTTGGCTT 1111
DB 524 CTAGATCGATCTAAGTTGACCGATCAGTTAGCTGATATAGCTACCAAGATTTGGCTT 583
QY 1112 GACCTTGATCGAGAGATCCATGTTTCATGTTTACCTGGAAATGATTTGTATATGCTGAATG 1171
DB 584 GACCTTGATCGAGAGATCCATGTTTCATGTTTACCTGGAAATGATTTGTATATGCTGAATG 643
QY 1172 AAATCTGAATCTGTTGAAGTTAGATTGAATCTGAACTGCACTGTCAATGTTAGATTGAATCTGA 1231
DB 644 AAATCTGAATCTGTTGAAGTTAGATTGAATCTGAACTGCACTGTCAATGTTAGATTGAATCTGA 703
QY 1232 ACACGTGTTAAGTTAGATCAAGAGTTTGTGTATAGATTCTTCGAAACCTTAGGATTTGTAGT 1291
DB 704 TAACTGTTTAAAGTTAGATCAAGAGTTTGTGTATAGATTCTTCGAAACCTTAGGATTTGTAGT 763
QY 1292 GTCGTACGTTGAAACAGAAAGCTATTCTGATTCATCAATCAGGGTTTATTGACTGTATTGAA 1351
DB 764 GTCGTACGTTGAAACAGAAAGCTATTCTGATTCATCAATCAGGGTTTATTGACTGTATTGAA 823
QY 1352 CTCCTTTTGTGTTTGGCTGAGCTCATAAA 1379
DB 824 CTCCTTTTGTGTTTGGCTGAGCTCATAAA 851

RESULT 14

US-10-425-114-14693

; Sequence 14693, Application US/10425114

; Publication No. US2004003488A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E.

; APPLICANT: Tabaska, Jack E.

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 14693

; LENGTH: 1468

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

; FEATURE:

; OTHER INFORMATION: Clone ID: L1B24-131-F9_FLI

US-10-425-114-14693

Query Match 10.0%; Score 144; DB 12; Length 1468;
Best Local Similarity 92.6%; Pred. No. 5.8e-24;
Matches 163; Conservative 0; Mismatches 10; Indels 3; Gaps 1;

QY 776 CTCACCTTTCATCAGCCGCTTTTGAATCTCCGCGACTTCAGAGAGAACAGAGAGAG 835
DB 1 CTCACCTTTCATCAGCCGCTTTTGAATCTCCGCGACTTCAGAGAGAACAGAGAGAG 60
QY 836 ACTAAGAGAGAAAGTAAGAGATATCCAGAGATTCATCTCCGTTTGAATCTCTCA 895
DB 61 ACTAAGAGAGAAAGTAAGAGATATCCAGAGATTCATCTCCGTTTGAATCTCTCA 120
QY 896 ATCTCATCT 951
DB 121 ATCTCA---TCT 173

RESULT 15

US-09-887-384A-14

; Sequence 14, Application US/09887384A

; Patent No. US2002004992A1

; GENERAL INFORMATION:

```

; APPLICANT: Hamilton, Carol
; TITLE OF INVENTION: MODIFIED PROMOTERS
; FILE REFERENCE: actin promoters
; CURRENT APPLICATION NUMBER: US/09/887,384A
; CURRENT FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 94
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: DNA fragments
; OTHER INFORMATION: for assembling an actin-based promoter
US-09-887-384A-14

Query Match      6.5%; Score 94; DB 9; Length 94;
Best Local Similarity 100.0%; Pred. No. 1.2e-12;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1059 CGATCTAAGTTGACCGATCAGTTAGCTCGATTATAGCTACCAAGATTGGCTTGACCTTG 1118
Db      1 CGATCTAAGTTGACCGATCAGTTAGCTCGATTATAGCTACCAAGATTGGCTTGACCTTG 60

QY      1119 ATGGAGATCCATGTTTCATGTTACCTGGGAAT 1152
Db      61 ATGGAGATCCATGTTTCATGTTACCTGGGAAT 94

```

Search completed: March 15, 2004, 05:20:22
Job time : 1422 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2004, 22:37:13 ; Search time 4096 Seconds

(without alignments)

10527.585 Million cell updates/sec

Title: US-09-868-744B-1

Perfect score: 1444

Sequence: 1 attatgatctcaataacatt.....gtactggaatgtagatccc 1444

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST: *

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estm:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_estl:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gssl:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
C 1	272.8	18.9	552	29	BX531590 Arabidops
2	202.4	14.0	327	9	AV831431
3	166	11.5	600	9	AV831862
4	163	11.3	546	9	AV525527

5	158	10.9	396	9	AV828177
6	158	10.9	562	9	AV830541
7	158	10.9	580	9	AV831261
8	158	10.9	667	9	AV824265
9	156	10.8	538	9	AV526249
10	153	10.6	399	9	AV829675
11	151	10.5	297	9	AV525898
12	150	10.4	162	28	BH901724
13	147	10.2	774	13	BU635286
14	146	10.1	463	14	N65062
15	144	10.0	438	14	N38049
16	144	10.0	597	14	T46735
17	143.8	10.0	382	9	AV545656
18	143	9.9	352	14	Z25952
19	143	9.9	393	14	T43759
20	142	9.8	411	14	R65270
21	139.4	9.7	253	14	T20889
22	138.4	9.6	897	14	CF652111
23	136	9.4	454	9	AV525074
24	136	9.4	512	9	AV525969
25	136	9.4	518	9	AV525842
26	132	9.1	370	14	T21821
27	132	9.1	446	14	CB253737
28	131	9.1	710	10	B8038458
29	131	9.1	915	14	CF652259
30	129.4	9.0	458	14	CB264030
31	126	8.7	540	9	AV550855
32	125	8.7	449	14	CF651398
33	124	8.6	412	14	R29797
34	123	8.5	426	9	AV554326
35	123	8.5	527	9	AV553948
36	123	8.5	546	9	AV518519
37	123	8.5	551	9	AV552230
38	120.6	8.4	251	14	CB258935
39	119.8	8.3	270	9	AV519680
40	119	8.2	491	14	T76037
41	119	8.2	502	14	N37237
42	118	8.2	311	14	T20714
43	118	8.2	478	14	R30501
44	116.2	8.0	293	9	AV528010
45	114.8	8.0	821	14	CF651397

ALIGNMENTS

RESULT 1	BX531590/c	552 bp	DNA	linear	GSS 03-JUN-2003
LOCUS	Arabidopsis thaliana T-DNA flanking sequence GK-481G03-019880,				
DEFINITION	genomic survey sequence.				
ACCESSION	BX531590				
VERSION	BX531590.1				GI:31408720
KEYWORDS	GSS.				
SOURCE	Arabidopsis thaliana (thale cress)				
ORGANISM	Arabidopsis thaliana				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.				
REFERENCE	1				
AUTHORS	Srithov,N., Li,Y., Rosso,M., Viechover,P., Dekker,K., Saedler,H. and Weisshaar,B.				
TITLE	A pipeline for automated high-throughput generation of FSTs (flanking sequence tags) from Arabidopsis thaliana T-DNA transformed lines				
JOURNAL	Unpublished				
REFERENCE	2				
AUTHORS	Rosso,M., Strizhov,N., Li,Y., Reiss,B., Dekker,K. and Weisshaar,B.				
TITLE	A new Arabidopsis thaliana T-DNA mutagenised population (GABI-kat) for flanking sequence tag based reverse genetics				
JOURNAL	Unpublished				
REFERENCE	3				
AUTHORS	Li,Y., Strizhov,N., Rosso,M. and Weisshaar,B.				

TITLE Direct Submission
JOURNAL Submitted (02-JUN-2003) Weishaar B., Max-Planck-Institut fuer Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT This sequence is recovered from the left border of the T-DNA. It indicates an insertion close to or within gene At3g18780. The sequences are generated at the MPI for Plant Breeding Research in the context of the GABI-kat project. GABI-kat is part of the German Plant Genomics program designated 'GABI'. Information on line availability can be found at: <http://www.mpiz-koeln.mpg.de/GABI-Kat/>.

FEATURES Location/Qualifiers
 1..552
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /strain="Columbia 0"
 /db_xref="taxon:3702"
 /clone="GK-481G03-019880"
 /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
 /note="PCR was performed on DNA from Arabidopsis thaliana plants (T1) which were transformed with the T-DNA from vector pAC161. The lines contain one or more T-DNA insertions. The DNA fragment(s) resulting from the PCR were directly sequenced to determine the genomic sequence flanking the insertion. Sequences displaying significant similarity to the A. thaliana nuclear genome sequence were processed for submission. T-DNA derived sequences were removed"

ORIGIN
 Query Match 18.9%; Score 272.8; DB 29; Length 552;
 Best Local Similarity 93.9%; Pred. No. 4.3e-50;
 Matches 294; Conservative 0; Mismatches 18; Indels 1; Gaps 1;

QY 1126 GATCCATGTTTCATGTTTACCTGGAAATGATTTGTATGTGAATCGAAATCGAAGCTTT 1185
 DB 552 GATCCATGTTTCATGTTTACCTGGAAATGATTTGTATGTGAATCGAAATCGAAGCTTT 493
 QY 1186 GAAGTTAGATGAATCGAATCGAATCGAATCGAATCGAATCGAATCGAATCGAATCGAAT 1244
 DB 492 GAAGTTAGATGAATCGAATCGAATCGAATCGAATCGAATCGAATCGAATCGAATCGAAT 433
 QY 1245 TAGATGAAGTTTCTGTATAGATTTCTCGAAACCTTAGGATTTGTAGTGTGCTGAGTGA 1304
 DB 432 TAGATGAAGTTTCTGTATAGATTTCTCGAAACCTTAGGATTTGTAGTGTGCTGAGTGA 373
 QY 1305 CAGAAAGCTATTTCTGATTCATCAATCAGGGTTTATTTGACTGTATGAACTCTTTTGTGTG 1364
 DB 372 CAGAAAGCTATTTCTGATTCATCAATCAGGGTTTATTTGCTGTATGAACTCTTTTGTGTG 313
 QY 1365 TTTCAGCTCATAAAAATGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 1424
 DB 312 TTTCAGCTCATAAAAATGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 253
 QY 1425 GTACTGGAATGGT 1437
 DB 252 GTACCGGTATGGT 240

RESULT 2
LOCUS AV831431
DEFINITION AV831431 RAF19 Arabidopsis thaliana cDNA clone RAF109-89-B08 5', mRNA sequence.
ACCESSION AV831431
VERSION AV831431.1 GI:19873491
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE 1 (bases 1 to 327)
AUTHORS Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M., Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Muramatsu, A., Hayashizaki, Y.

Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y. Large scale analysis of Arabidopsis full-length cDNA (2002b) Unpublished (2002)
COMMENT Contact: Motoaki Seki
 Plant Functional Genomics Research Group
 RIKEN Genomic Sciences Center
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
 Tel: 81-298-36-4359
 Fax: 81-298-36-9060
 Email: maeki@rkc.riken.go.jp
 An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XhoI was ligated to modified lambdaBda F1C-1 vector (Carninci et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified pBluescript vector. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further details.

FEATURES Location/Qualifiers
 1..327
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /clone="RAF109-89-B08"
 /dev_stage="plants at various developmental stages from germination to mature seeds"
 /clone_lib="RAFL9"
 /lab_host="DH10B"
 /notes="Site 1: BamHI; Site 2: SalI; subjected to dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24 hr) treatments"

ORIGIN
 Query Match 14.0%; Score 202.4; DB 9; Length 327;
 Best Local Similarity 94.1%; Pred. No. 1.9e-34;
 Matches 222; Conservative 0; Mismatches 11; Indels 3; Gaps 1;
 QY 716 AAACCCGCTATATAATTCATATATTTCTCCCGCTTTGATTTGCTGCTGCTC 775
 DB 3 AAACCCGCTATATAATTCATATATTTCTCCCGCTTTGATTTGCTGCTGCTC 62
 QY 776 CTCATCTTCATCAGCGGTTTGAATCTCCGCGACTTGACAGAGAACAAGGAAG 835
 DB 63 CTCATCTTCATCAGCGGTTTGAATCTCCGCGACTTGACAGAGAACAAGGAAG 122
 QY 836 ACTAAGAGAGAAAGTAAGAGATAATCCAGAGATTCATCTCCGTTTGAATCTTCCTCA 895
 DB 123 ACTAAGAGAGAAAGTAAGAGATAATCCAGAGATTCATCTCCGTTTGAATCTTCCTCA 182
 QY 896 ATCTCATCTTCTCTCCGCTTTCTTTTCCAGGTAATAGAACTTCTGGATCT 951
 DB 183 ATCTCA---TCTTCTCCGCTTTCTTTTCCAGGTAATAGAACTTCTGGATCT 235

RESULT 3
LOCUS AV831862
DEFINITION AV831862 RAF19 Arabidopsis thaliana cDNA clone RAF109-96-J07 5', mRNA sequence.
ACCESSION AV831862
VERSION AV831862.1 GI:19873922
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE 1 (bases 1 to 600)
AUTHORS Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M., Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Muramatsu, A., Hayashizaki, Y. and Shinozaki, K.

```
1. 546
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="APD25g08R"
/tissue_type="aboveground organs"
/dev_stage="two to six-week old"
/clone_lib="Arabidopsis thaliana
six-week old"
```

```

/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"

```

QY	617	GAGGAGCTCAGCGAATGAGAGACCTAAGGAGAAAGTAAAGAGTAATCCAGGAGAATTCAATCT
Db	61	GAGAGAACAAGGAAGAGACTAAGGAGAAAGTAAAGAGTAATCCAGGAGAATTCAATCT
QY	877	CGGTTTTGAACTTCCTCAATCTCATCTTCTTCTTCGGCTCTTCTTCCAAAGSTAATAG
Db	121	CGTTTTTGAACTTCCTCAATCTCA---TCITCTTCGGCTCTTCTTCCAAAGTCATAA
QY	937	GAACTTTCTGGATCT
Db	178	AAAATGGCTGAGGCT

REFERENCE
AUTHORS
Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M., Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y. and Shinozaki, K.
1 (bases 1 to 396)
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Brassicales; Brassicaceae; Arabidopsids.
An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XhoI was ligated to modified Lambda FIC-1 vector (Carninci et al., submitted for publication) digested with BamHI and XhoI. This clone is in a modified Bluescript vector. Please visit our web
TITLE
Large scale analysis of Arabidopsis full-length cDNA (2002b)
JOURNAL
Unpublished (2002)
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Tel: 81-298-36-9060
Fax: 81-298-36-9060
Email: meeki@etc.riken.go.jp

site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further details.

FEATURES

source

Location/Qualifiers

```
1..396
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL09-24-N21"
/dev_stage="plants at various developmental stages from
germination to mature seeds"
/lab_host="DH10B"
/clone_lib="RAFL9"
/notes="Site 1: BamHI; Site 2: SalI; subjected to
dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24
hr) treatments"
```

ORIGIN

```
Query Match      10.9%; Score 158; DB 9; Length 396;
Best Local Similarity 93.2%; Pred. No. 1.3e-24;
Matches 177; Conservative 0; Mismatches 10; Indels 3; Gaps 1;

QY 762 GTCTCGTGTCTCTCTCACTTTCATCAGCCGTTTGAATCTCGGCGACTTGACAGAA 821
Db 2 GTCTCGTGTCTCTCTCACTTTCATCAGCCGTTTGAATCTCGGCGACTTGACAGAA 61

QY 822 GAACAAGGAGAGACTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCTTCGTT 881
Db 62 GAACAAGGAGAGACTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCTTCGTT 121

QY 882 TTGAATCTTCCTCAATCTCATCTCTTCTTCGCTCTTTTCTTCCAAGGTAATAGAACT 941
Db 122 TTGAATCTTCCTCAATCTCA---TCTTCTTCGCTCTTTTCTTCCAAGGTAATAGAACT 178

QY 942 TTCTGGATCT 951
Db 179 GGCTGAGGCT 188
```

RESULT 6

```
AV830541
LOCUS AV830541 RAFL9 Arabidopsis thaliana cDNA clone RAFL09-70-008 5',
DEFINITION mRNA sequence.
ACCESSION AV830541
VERSION AV830541.1 GI:19872601
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 562)
Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,
Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,
Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.
and Shinozaki, K.
Large scale analysis of Arabidopsis full-length cDNA (2002b)
Unpublished (2002)
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: meeki@rc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
and XhoI was ligated to modified Lambda FIC-1 vector (Carninci et
al., submitted for publication) digested with BamHI and SalI. This
clone is in a modified pBluescript vector. Please visit our web
site (http://www.gsc.riken.go.jp/e/plant/index\_e.html) for further
details.
```

TITLE

JOURNAL

COMMENT

Unpublished (2002)

Contact: Motoaki Seki

Plant Functional Genomics Research Group

RIKEN Genomic Sciences Center

3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan

Tel: 81-298-36-4359

Fax: 81-298-36-9060

Email: meeki@rc.riken.go.jp

An Arabidopsis full-length cDNA library was constructed essentially

as reported previously (Seki et al., 1998). cDNA cleaved with BamHI

and XhoI was ligated to modified Lambda FIC-1 vector (Carninci et

al., submitted for publication) digested with BamHI and SalI. This

clone is in a modified pBluescript vector. Please visit our web

site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further

details.

FEATURES

source

Location/Qualifiers

```
1..580
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
```

source

1..562

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/clone="RAFL09-70-008"

/dev_stage="plants at various developmental stages from

germination to mature seeds"

/lab_host="DH10B"

/clone_lib="RAFL9"

/notes="Site 1: BamHI; Site 2: SalI; subjected to

dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24

hr) treatments"

ORIGIN

```
Query Match      10.9%; Score 158; DB 9; Length 562;
Best Local Similarity 93.2%; Pred. No. 1.3e-24;
Matches 177; Conservative 0; Mismatches 10; Indels 3; Gaps 1;

QY 762 GTCTCGTGTCTCTCTCACTTTCATCAGCCGTTTGAATCTCGGCGACTTGACAGAA 821
Db 2 GTCTCGTGTCTCTCTCACTTTCATCAGCCGTTTGAATCTCGGCGACTTGACAGAA 61

QY 822 GAACAAGGAGAGACTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCTTCGTT 881
Db 62 GAACAAGGAGAGACTAAGAGAGAAAGTAAGAGATAATCCAGGAGATTCTTCGTT 121

QY 882 TTGAATCTTCCTCAATCTCATCTCTTCTTCGCTCTTTTCTTCCAAGGTAATAGAACT 941
Db 122 TTGAATCTTCCTCAATCTCA---TCTTCTTCGCTCTTTTCTTCCAAGGTAATAGAACT 178

QY 942 TTCTGGATCT 951
Db 179 GGCTGAGGCT 188
```

RESULT 7

```
AV831261
LOCUS AV831261 RAFL9 Arabidopsis thaliana cDNA clone RAFL09-87-G17 5',
DEFINITION mRNA sequence.
ACCESSION AV831261
VERSION AV831261.1 GI:19873321
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 580)
Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,
Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,
Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.
and Shinozaki, K.
Large scale analysis of Arabidopsis full-length cDNA (2002b)
Unpublished (2002)
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: meeki@rc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
and XhoI was ligated to modified Lambda FIC-1 vector (Carninci et
al., submitted for publication) digested with BamHI and SalI. This
clone is in a modified pBluescript vector. Please visit our web
site (http://www.gsc.riken.go.jp/e/plant/index\_e.html) for further
details.
```

TITLE

JOURNAL

COMMENT

Unpublished (2002)

Contact: Motoaki Seki

Plant Functional Genomics Research Group

RIKEN Genomic Sciences Center

3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan

Tel: 81-298-36-4359

Fax: 81-298-36-9060

Email: meeki@rc.riken.go.jp

An Arabidopsis full-length cDNA library was constructed essentially

as reported previously (Seki et al., 1998). cDNA cleaved with BamHI

and XhoI was ligated to modified Lambda FIC-1 vector (Carninci et

al., submitted for publication) digested with BamHI and SalI. This

clone is in a modified pBluescript vector. Please visit our web

site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further

details.

FEATURES

source

Location/Qualifiers

```
1..580
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
```

/db_xref="taxon:3702"
 /clone="RAFL09-87-G17"
 /dev_stage="plants at various developmental stages from germination to mature seeds"
 /lab_host="DH10B"
 /clone_lib="RAFL9"
 /note="Site 1: BamHI; Site 2: SalI; subjected to dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24 hr) treatments"

ORIGIN

Query Match 10.9%; Score 158; DB 9; Length 580;
 Best Local Similarity 93.2%; Pred. No. 1.3e-24;
 Matches 177; Conservative 0; Mismatches 10; Indels 3; Gaps 1;
 QY 762 GTCTGTTGTCCTCTCACTTTTCATCAGCCGTTTGAATCTCCGGGACTTCACAGAGAA 821
 Db 1 GTCTGTTGTCCTCTCACTTTTCATCAGCCGTTTGAATCTCCGGGACTTCACAGAGAA 60
 QY 822 GAACAGAGAGAGACTTAAGAGAGAAAGTAAGAGATAATCCAGAGATTCATTCCTCGTT 881
 Db 61 GAACAGAGAGAGACTTAAGAGAGAAAGTAAGAGATAATCCAGAGATTCATTCCTCGTT 120
 QY 882 TTGAATCTTCTCAATCTCATCTTCTTCCTCCGCTCTTTTCCAGGTAATAGGAAT 941
 Db 121 TTGAATCTTCTCAATCTCA--TCTTCTCCGCTCTTTTCCAGGTAATAGGAAT 177
 QY 942 TTCTGGATCT 951
 Db 178 GCGTGAGGCT 187

RESULT 8

AV824265
 LOCUS AV824265 RAFL6 Arabidopsis thaliana cDNA clone RAFL06-16-M11 5', mRNA sequence.
 DEFINITION

ACCESSION AV824265

VERSION AV824265.1 GI:19866325

KEYWORDS EST.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 667)

Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M., Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y. and Shinozaki, K.

Large scale analysis of Arabidopsis full-length cDNA (2002b)

Unpublished (2002)

Contact: Motoaki Seki

Plant Functional Genomics Research Group

RIKEN Genomic Sciences Center

3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan

Tel: 81-298-36-4359

Fax: 81-298-36-9060

Email: mseki@rtc.riken.go.jp

An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). This clone is in a modified pBluescript vector as a SstI/XhoI insert. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further details.

Location/Qualifiers

1..667

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/clone="RAFL06-16-M11"

/dev_stage="plants at various developmental stages from germination to mature seeds"

/lab_host="DH10B"

FEATURES

source

/clone_lib="RAFL6"
 /note="Site 1: SstI; Site 2: XhoI; subjected to dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24 hr) treatments"

ORIGIN
 Query Match 10.9%; Score 158; DB 9; Length 667;
 Best Local Similarity 93.2%; Pred. No. 1.2e-24;
 Matches 177; Conservative 0; Mismatches 10; Indels 3; Gaps 1;
 QY 762 GTCTGTTGTCCTCTCACTTTTCATCAGCCGTTTGAATCTCCGGGACTTCACAGAGAA 821
 Db 2 GTCTGTTGTCCTCTCACTTTTCATCAGCCGTTTGAATCTCCGGGACTTCACAGAGAA 61
 QY 822 GAACAGAGAGAGACTTAAGAGAGAAAGTAAGAGATAATCCAGAGATTCATTCCTCGTT 881
 Db 62 GAACAGAGAGAGACTTAAGAGAGAAAGTAAGAGATAATCCAGAGATTCATTCCTCGTT 121
 QY 882 TTGAATCTTCTCAATCTCATCTTCTTCCTCCGCTCTTTTCCAGGTAATAGGAAT 941
 Db 122 TTGAATCTTCTCAATCTCA--TCTTCTCCGCTCTTTTCCAGGTAATAGGAAT 178
 QY 942 TTCTGGATCT 951
 Db 179 GCGTGAGGCT 188

RESULT 9

AV526249
 LOCUS AV526249 538 bp mRNA linear EST 01-SEP-2000
 DEFINITION Old Arabidopsis thaliana aboveground organs two to six-week old Arabidopsis thaliana cDNA clone APZ08b09R 5', mRNA sequence.

ACCESSION AV526249

VERSION AV526249.1 GI:8685777

KEYWORDS EST.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 538)

Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.

A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries

DNA Res. 7 (3), 175-180 (2000)

20363093

10907847

PUBMED

COMMENT

Contact: Erika Asamizu

The First Laboratory for Plant Gene Research

Kazusa DNA Research Institute

Yana 1532-3, Kisarazu, Chiba 292-0812, Japan

Email: asamizu@kazusa.or.jp; URL: <http://www.kazusa.or.jp/en/plant/>.

Location/Qualifiers

1..538

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/strain="Columbia"

/db_xref="taxon:3702"

/clone="APZ08b09R"

/tissue_type="aboveground organs"

/dev_stage="two to six-week old"

/clone_lib="Arabidopsis thaliana aboveground organs two to six-week old"

/note="vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match 10.8%; Score 156; DB 9; Length 538;
 Best Local Similarity 93.1%; Pred. No. 3.5e-24;
 Matches 175; Conservative 0; Mismatches 10; Indels 3; Gaps 1;
 QY 764 CTGTTGTCCTCTCACTTTTCATCAGCCGTTTGAATCTCCGGGACTTCACAGAGAA 823

827	AGGAGAGAGACTAAGAGAGAAAGTAAGAGAGTAATCCAGGAGATTCAATCTCCGTTTTGAA	88
QY		
62	AGGAGAGAGCTAAGAGAGAAAGTAAGAGATTAATCCAGGAGATTCAATCTCCGTTTTGAA	121
Db		
887	TCTTCCCTCAATCATCTATCTTCTTCGCGCTCTTCTTTCCAAGGTAATAGGAACATTCTCTG	946
QY		
122	TCTTCCCTCAATCTCA---TCTTCTCCGCGCTCTTCTTTCCAAGCTCATAAAATGGCTG	178
Db		
947	GATCT	951
QY		
179	AGGCT	183
Db		

RESULT 11	AV525898	297 bp	linear	EST 01-SEP-2000
LOCUS	AV525898			
DEFINITION	Arabidopsis thaliana aboveground organs two to six-week old Arabidopsis thaliana cDNA clone AP332806R 5', mRNA sequence.			

SOURCE	ORGANISM	PREFERENCE
Arabiopsis thaliana (Cruciferae)	Arabiopsis thaliana	1 (bases 1 to 297)
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;	
	rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.	

COMMENT

Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 232-0812, Japan
Email: asamizu@kazusa.or.jp, URL: <http://www.kazusa.or.jp/en/plant/>.

```

ORIGIN
      Query Match      10.0%; Score 151; DB 9; Length 297;
      Best Local Similarity 92.0%; Pred. No. 4.8e-23;
      Matches 170; Conservative 0; Mismatches 10; Indels 3; Caps 1;
      /note="vector: paluescriptII SK-; Site_1: EcoRI; Site_2:
      XhoI"

```

Db	61	GAAGAAGACTAAGAGAGAAAGTAAGAGATATCCAGGAGATTCAATTCTCCGTTTGGATC	120
Qy	889	TTCTCTCAATCTCATCTCTTCTTTCCTTCGCTCTTCTTTCCTTCGAGTAAATAGAGACTTCTTCGA	948
Db	121	TTCTCTCAATCTCA---TCTTCTTCGCTCTCTTCTTTCCTTCGAGCTCATAAAAAATGGCTGAG	177
Qy	949	TCT	951
Db	178	GCT	180

COMMENT

Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
Lansing, MI

Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313tcn@ibm.cl.msu.edu
Seq primer: T7 dye primer.

FEATURES

source

Location/Qualifiers

1..463

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/strain="var columbia"

/db_xref="taxon:3702"

/clone="226A8T7"

/clone_lib="Lambda-PRL2"

/note="Vector: lambda Zip-Lox; Site 1: Sal; Site 2: Not; Lambda PRL2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques). The vector is BRL's lambda Zip-Lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dT primed cDNA. "

ORIGIN

Query Match 10.1%; Score 146; DB 14; Length 463;

Best local Similarity 92.1%; Pred. No. 6e-22; Mismatches 10; Indels 3; Gaps 1;

Matches 165; Conservative 0; Mismatches 10; Indels 3; Gaps 1;

QY 774 TCCTCCTTCATCAGCGGTTTGAATCTCCGGCAGCTTGACAGAGAAGCAAGGAAGA 833
DB 1 TCCTCCTTCATCAGCGGTTTGAATCTCCGGCAGCTTGACAGAGAAGCAAGGAAGA 60
QY 834 AGACTAAGAGAAGTAAGAGATAATCCAGGAGATTCATTCCTCGTTTGAATCTTCT 893
DB 61 AGACTAAGAGAAGTAAGAGATAATCCAGGAGATTCATTCCTCGTTTGAATCTTCT 120
QY 894 CAATCTCATCTTCTTCCTCGCTCTTTTCCAAAGTAAATAGGAACCTTCTGGATCT 951
DB 121 CAATCTCA---TCCTTCCTCGCTCTTTTCCAAAGTAAATAGGAACCTTCTGGATCT 175

RESULT 15

N38049

LOCUS

N38049 13276 Lambda-PRL2 Arabidopsis thaliana cDNA clone 217C7T7, mRNA

sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 438)

REFERENCE

AUTHORS

McIntosh, L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomashow, M.,

Retzel, E. and Somerville, C.

Genes galore: a summary of methods for accessing results from

large-scale partial sequencing of anonymous Arabidopsis cDNA clones

Plant Physiol. 106, 1241-1255 (1994)

JOURNAL

MEDLINE

PUBMED

7846151

COMMENT

Contact: Thomas Newman

MSU-DOE Plant Research Laboratory

Michigan State University

MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.

Lansing, MI

Tel: 517-353-0854

Fax: 517-353-9168

Email: 22313tcn@ibm.cl.msu.edu

Seq primer: T7 dye primer.

FEATURES

source

Location/Qualifiers

1..438

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/strain="var columbia"

/db_xref="taxon:3702"

/clone="217C7T7"

/clone_lib="Lambda-PRL2"

/note="Vector: lambda Zip-Lox; Site 1: Sal; Site 2: Not; Lambda PRL2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques). The vector is BRL's lambda Zip-Lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dT primed cDNA. "

ORIGIN

Query Match 10.0%; Score 144; DB 14; Length 438;

Best local Similarity 92.6%; Pred. No. 1.7e-21; Mismatches 10; Indels 3; Gaps 1;

Matches 163; Conservative 0; Mismatches 10; Indels 3; Gaps 1;

QY 776 CTCACCTTTCATCAGCGGTTTGAATCTCCGGCAGCTTGACAGAGAAGCAAGGAAGAAG 835
DB 1 CTCACCTTTCATCAGCGGTTTGAATCTCCGGCAGCTTGACAGAGAAGCAAGGAAGAAG 60
QY 836 ACTAAGAGAAGTAAGAGATAATCCAGGAGATTCATTCCTCGTTTGAATCTTCTCTCA 895
DB 61 ACTAAGAGAAGTAAGAGATAATCCAGGAGATTCATTCCTCGTTTGAATCTTCTCTCA 120
QY 896 ATCTCATCTTCTTCTCGCTCTTTTCCAAAGTAAATAGGAACCTTCTGGATCT 951
DB 121 ATCTCA---TCCTTCCTCGCTCTTTTCCAAAGTAAATAGGAACCTTCTGGATCT 173

Search completed: March 15, 2004, 01:43:50

Job time : 4103 secs